



**PHD**

**Personal Informatics Systems and the Integration of Data from Novel Sensor Technologies**

Dudley, Cillian

*Award date:*  
2021

*Awarding institution:*  
University of Bath

[Link to publication](#)

**Alternative formats**

If you require this document in an alternative format, please contact:  
[openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk)

Copyright of this thesis rests with the author. Access is subject to the above licence, if given. If no licence is specified above, original content in this thesis is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International (CC BY-NC-ND 4.0) Licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>). Any third-party copyright material present remains the property of its respective owner(s) and is licensed under its existing terms.

**Take down policy**

If you consider content within Bath's Research Portal to be in breach of UK law, please contact: [openaccess@bath.ac.uk](mailto:openaccess@bath.ac.uk) with the details. Your claim will be investigated and, where appropriate, the item will be removed from public view as soon as possible.

# Personal Informatics Systems and the Integration of Data from Novel Sensor Technologies

submitted by

Cillian Dudley

for the degree of Doctor of Philosophy

of the

University of Bath

Department of Computer Science

March, 2021

## **COPYRIGHT**

Attention is drawn to the fact that copyright of this thesis rests with the author. A copy of this thesis has been supplied on condition that anyone who consults it is understood to recognise that its copyright rests with the author and that they must not copy it or use material from it except as permitted by law or with the consent of the author.

This thesis may be made available for consultation  
within the University Library and may be  
photocopied or lent to other libraries for the purposes  
of consultation with effect from.....(date)

Signed on behalf of the Faculty of Science .....



## Abstract

Personal Informatics (PI) systems are used by people to record, track and analyse data about themselves. The increasing availability and affordability of sensing technologies, which is likely to continue into the future, provides people with the ability to record and analyse an ever increasing amount of data about themselves. Primarily, research relating to Personal Informatics seeks to understand how people are using currently available, widely adopted sensing technologies, or how specific, mainstream sensing technologies might be used to solve a specific problem. Currently, there is little understanding of how people attempt to make sense of data from newly emerging sensing technologies; those which do not have widespread adoption, or whose data is less understood by the general population. Given the rapid expansion of sensing capabilities, the research described in this thesis provides an important understanding of how people use emerging sensing technologies to learn more about themselves, as well as the challenges and opportunities these technology present in relation to self-understanding. Throughout our research we used a NeuroSky brain-computer interface as a technology probe, representing a novel sensing technology, and developed a feature-rich multifaceted PI system for allowing users to engage with data. Our findings stem from qualitative analysis of participant interviews in one exploratory study and two in-the-wild studies of these novel sensing technologies. Additionally, we explore aspects of trust in novel sensing technologies. We present quantitative and qualitative analysis of participant responses from a lab-based study designed to explore how perceptions of trust are shaped. We present several design considerations and challenges for developing PI systems that integrate novel sensing technologies. These considerations and challenges are based around two groups of users with distinct tracking behaviours: those tracking for documentary purposes - to gather data that documents their lives - and those with a specific life-aspect that motivates data tracking, such as those living with chronic health conditions or symptoms. Specifically, we explore the impact of novel sensing technologies for those with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS), prolonged fatigue, chronic fatigue (CF), and idiopathic chronic fatigue (ICF). We additionally present four design considerations for the development of features which enable users to interrogate their data, using automatically generated insights and predefined analyses. We highlight that trust in devices may be shaped by users' preconceived initial levels of trust in a device, rather than on specific feedback from the device or visualisations of the data that it generates.



# Acknowledgements

I would like to thank my supervisor, Dr. Simon L. Jones, for his continued guidance, support and advice during the course of my PhD. I am deeply grateful to all the participants who took part in this research, without whom none of this would have been possible.

I would also like to thank my family, Noeleen, Séamus, Ciara and Alfie, for their support throughout my degree. Also, my friends for putting up with me and providing much needed distractions from PhD life: Jamie, Meghna, David, and Thomas.

Finally, everyone in the Department of Computer Science and all the people I've met through conferences and summer schools, for being amazing people.

# Table of Contents

<b>List of Figures</b>	<b>10</b>
<b>List of Tables</b>	<b>12</b>
<b>List of Abbreviations</b>	<b>14</b>
<b>1 Introduction</b>	<b>16</b>
1.1 Personal Informatics . . . . .	17
1.2 Motivation for this Research . . . . .	21
1.3 Thesis Structure & Contributions . . . . .	24
<b>2 Literature Review</b>	<b>28</b>
2.1 Chapter Overview . . . . .	29
2.2 Personal Informatics . . . . .	29
2.2.1 Personal Informatics Research . . . . .	30
2.2.2 Understanding Tracking Practices . . . . .	31
2.2.3 Personal Informatics Models . . . . .	32
2.2.3.1 Stage-Based Model . . . . .	33
2.2.3.2 Lived-Informatics Model . . . . .	34
2.2.4 Recommendations for Design Features . . . . .	35
2.2.5 Visualisation Techniques . . . . .	36
2.2.6 Personal Informatics Systems . . . . .	37
2.3 Novel Sensor Technologies . . . . .	39
2.3.1 Data & Sensor Availability Increases, Cost Decreases . . . . .	40
2.3.2 Sensors are Wearable . . . . .	41
2.3.3 Data Collection is Automated . . . . .	44
2.3.4 Data Collection is Unobtrusive . . . . .	45
2.3.5 Data Collection & Analysis is Personal . . . . .	45

2.3.6	Data is Multifaceted . . . . .	46
2.3.7	Transparency of Data Collection is Unclear . . . . .	46
2.3.8	Data May Have Varying Levels Trustworthiness . . . . .	48
2.3.9	Data May Lack Familiarity . . . . .	49
2.3.10	Data has Volume, Variety, Velocity & Veracity . . . . .	50
2.4	Sensors Used in This Research . . . . .	51
2.4.1	Heart Rate Sensor - Fitbit . . . . .	52
2.4.2	Breathing Rate Sensor - Spire . . . . .	52
2.4.3	Electroencephalography - NeuroSky MindWave . . . . .	53
2.5	Technology Probe Sensor Choice . . . . .	54
2.6	Electroencephalography and Brain-Computer Interfaces . . . . .	55
2.6.1	Electroencephalography . . . . .	56
2.6.2	Brain-Computer Interfaces . . . . .	56
2.6.3	Consumer-Grade BCI . . . . .	58
2.7	Brain-Computer Interfaces for Personal Informatics . . . . .	59
2.7.1	Fitbit for the Brain . . . . .	61
2.7.2	Commercially available technologies . . . . .	61
2.7.3	Neurotechnology & Neuroergonomics . . . . .	62
2.8	Trust in PI . . . . .	63
2.9	Chapter Summary . . . . .	66
<b>3</b>	<b>Fitbit for the Mind?: An Exploratory Study of Personal Informatics</b>	<b>67</b>
3.1	Chapter Overview . . . . .	68
3.2	Methodology . . . . .	69
3.2.1	Cognition Tracker App . . . . .	69
3.2.2	Participants . . . . .	71
3.2.3	Procedure . . . . .	72
3.3	Interview Results . . . . .	73
3.3.1	Why Use Cognitive Personal Informatics Systems? . . . . .	73
3.3.1.1	CPI for Monitoring Medical Conditions . . . . .	73
3.3.1.2	CPI for Monitoring General Well-being . . . . .	74
3.3.1.3	Improving Understanding of Self & Optimising Behaviour	74
3.3.1.4	Improving Understanding of Others . . . . .	75
3.3.1.5	Supplementing Existing Tracking Tools . . . . .	75
3.3.1.6	Hobbyist/Technophile Uses & Hedonic Experience . . . . .	75
3.3.2	What Insights Will Cognitive Personal Informatics Systems Pro- vide? . . . . .	76

3.3.2.1	Analysing EEG Data in a CPI System . . . . .	77
3.3.2.2	Data Integration . . . . .	79
3.3.3	Barriers to Use for Cognitive Personal Informatics . . . . .	79
3.3.3.1	Ergonomic & Aesthetic Hardware Issues . . . . .	79
3.3.3.2	Difficulties Interpreting EEG Data . . . . .	81
3.3.3.3	Privacy and Security Concerns . . . . .	82
3.3.3.4	Health Concerns . . . . .	82
3.3.3.5	Interference of Real-Time Feedback . . . . .	83
3.4	Discussion . . . . .	83
3.4.1	Challenge 1: Addressing the indirection between meaningful psychological states and brain activity data . . . . .	84
3.4.2	Challenge 2: Supporting diverse tracking styles . . . . .	85
3.4.3	Challenge 3: Encouraging exploratory & enquiring approaches . . . . .	86
3.4.4	Challenge 4: Overcoming misconceptions & lack of understanding about the brain . . . . .	87
3.5	Conclusion . . . . .	87
3.6	Addressing These Challenges . . . . .	88
<b>4</b>	<b>How People Use a Novel Sensor PI System</b>	<b>90</b>
4.1	Chapter Overview . . . . .	91
4.2	Methodology . . . . .	92
4.2.1	Mobile Application . . . . .	92
4.2.1.1	Session Recording . . . . .	94
4.2.1.2	Session Review . . . . .	94
4.2.2	Participants . . . . .	98
4.2.3	Procedure . . . . .	99
4.2.4	Interviews . . . . .	99
4.3	Interview Results . . . . .	100
4.3.1	How do users record, collect, analyse and make sense of data from ‘novel sensor technologies’ using a multifaceted PI system? . . . . .	101
4.3.1.1	Willingness to Record with Novel Sensing Technologies . . . . .	101
4.3.1.2	Novel Sensing Technology Concerns . . . . .	104
4.3.1.3	Scheduling and Timing . . . . .	105
4.3.1.4	Variation of Tracking Contexts . . . . .	106
4.3.2	What difficulties do users have in making sense of sensor data? . . . . .	106
4.3.2.1	Insufficient Knowledge or Information . . . . .	107
4.3.2.2	Amount of Data Shown . . . . .	108

4.3.2.3	Gaining Insight from Data . . . . .	108
4.3.2.4	Effect of Live View During Recording on Analysis . . .	109
4.3.2.5	Preferring Specific Types and Sources of Data . . . . .	110
4.3.3	How do users appropriate the data from the tracking application - what visualisations, records, additional analyses etc. do they create and why? . . . . .	111
4.3.3.1	Analysing Data . . . . .	111
4.3.4	What are the implications of providing people with metrics which have not yet been clinically validated? . . . . .	114
4.3.4.1	Trust . . . . .	115
4.3.4.2	Accuracy . . . . .	116
4.3.5	Does the use of abstracted metric values rather than wavebands have an effect on self-understanding/self-reflection/self-improvement? . . . . .	117
4.3.5.1	Difficulty Interpreting Metrics . . . . .	117
4.3.5.2	Data Meeting Expectations . . . . .	118
4.4	Discussion . . . . .	118
4.4.1	Discreet & Unobtrusive to Everyday Life . . . . .	118
4.4.2	Knowledge Transfer . . . . .	119
4.4.3	Simplify Insight Extraction . . . . .	119
4.4.4	Build Trust in Novel Sensing Technologies . . . . .	120
4.4.5	Assessing Accuracy in Novel Sensing Technologies . . . . .	120
4.5	Limitations . . . . .	121
4.6	Conclusion . . . . .	121
<b>5</b>	<b>Making Sense of Novel Sensor Data Within The Context of Fatigue</b>	<b>123</b>
5.1	Chapter Overview . . . . .	124
5.2	Methodology . . . . .	126
5.2.1	Application . . . . .	126
5.2.2	Participants . . . . .	132
5.3	Interview Results . . . . .	133
5.3.1	How do people with ME/CFS, CF or ICF perceive benefits and limitations of data from novel sensing technologies? . . . . .	133
5.3.1.1	Immediacy of Data . . . . .	133
5.3.1.2	Difficulties Due to ME/CFS, CF or ICF . . . . .	136
5.3.1.3	Tracking Data of Interest . . . . .	137

5.3.2	How do people living with ME/CFS, CF or ICF attempt to make sense of data from a range of novel sensors, for which they have varying degrees of understanding? What problems do the encounter? . . . . .	138
5.3.2.1	Target Setting . . . . .	139
5.3.2.2	Comparisons . . . . .	140
5.3.2.3	Data Not Meeting Expectations . . . . .	141
5.3.2.4	Insufficient Knowledge or Information . . . . .	142
5.3.2.5	Data Presentation . . . . .	144
5.3.3	How do people determine if a device is trustworthy? . . . . .	147
5.3.3.1	Own Research . . . . .	147
5.3.3.2	Compare and Correlate Sensor Results . . . . .	147
5.3.3.3	Inherent Trust in Technology . . . . .	148
5.3.3.4	Trust Values That Meet Expectations . . . . .	148
5.3.3.5	Trust Values That Are More Accurate . . . . .	148
5.4	Discussion . . . . .	149
5.4.1	Designs That Consider The Effects A Condition May Have on Users . . . . .	149
5.4.2	Supporting longer-term tracking . . . . .	150
5.4.3	Supporting target-setting and comparisons when previous data may not be available . . . . .	152
5.4.4	Supporting a variety of data presentation formats, varying levels of data knowledge, and technical knowledge . . . . .	153
5.4.5	Providing information about the basis on which the data is formed	154
5.5	Limitations . . . . .	154
5.6	Conclusion . . . . .	154
<b>6</b>	<b>Encouraging Exploratory &amp; Enquiring Approaches: Comparing the Insight Feed and Query Area Interfaces</b>	<b>156</b>
6.1	Chapter Overview . . . . .	157
6.2	Insight Feed . . . . .	159
6.2.1	Correlations . . . . .	163
6.2.2	Trends . . . . .	164
6.2.3	Extremes . . . . .	165
6.2.4	Variability . . . . .	165
6.3	Query Area . . . . .	166
6.4	Methodology . . . . .	166

6.5	Results . . . . .	168
6.5.1	Feature Usage Comparison . . . . .	168
6.5.2	Using the Insight Feed . . . . .	170
6.5.3	Utility of Different Insight Types . . . . .	171
6.5.4	Challenges with the Insight Feed . . . . .	177
6.5.5	Using the Query Area . . . . .	180
6.5.6	Utility of Queries . . . . .	181
6.5.7	Challenges Using the Query Area . . . . .	183
6.5.8	Time to Sense-making . . . . .	184
6.6	Discussion . . . . .	187
6.6.1	Designing for Exploration and Enquiry . . . . .	187
6.6.2	Designing to Uncover Insights - Volume & Variety . . . . .	189
6.6.3	Designing for the Sceptic . . . . .	192
6.6.4	Designing for Context-based Analysis . . . . .	192
6.7	Conclusion . . . . .	193
<b>7</b>	<b>Trust in Novel Sensing Technologies</b>	<b>195</b>
7.1	Chapter Overview . . . . .	196
7.2	Methodology . . . . .	199
7.2.1	Survey & Stroop Test . . . . .	201
7.2.2	Visualisation Manipulation . . . . .	204
7.2.3	Participant Deception Considerations . . . . .	209
7.2.4	Participants . . . . .	209
7.3	Results . . . . .	210
7.3.1	Confidence in Accuracy of Data Visualisation . . . . .	212
7.3.2	Match Between Visualised Data and Prior Expectations . . . . .	215
7.3.3	Exploratory Analysis . . . . .	216
7.3.4	Open-ended Response Analysis . . . . .	219
7.4	Discussion . . . . .	229
7.4.1	There will be a significant increase in user trust in data when the data more closely reflects users' prior predictions about data values. . . . .	229
7.4.2	There will be a significant difference in user trust in a personal informatics device and the data it produces according to the pre- cision at which it is presented within a visualisation. Therefore rounded values will be less trusted than precise values. . . . .	230

7.4.3	Trust will differ according to whether the data is presented synchronously (in real-time) vs. asynchronously (post-task), based on participants' ability to check 'in the moment' whether the data they are seeing corroborates their expectations. . . . .	231
7.4.4	Trust in the sensor data will differ based on the participants level of knowledge they have about how a particular sensing technology is measuring and capturing data. . . . .	231
7.4.5	Trust is Influenced by Initial Confidence . . . . .	231
7.5	Limitations . . . . .	232
7.6	Conclusion . . . . .	233
<b>8</b>	<b>Conclusion</b>	<b>234</b>
8.1	Thesis Summary . . . . .	235
8.2	Contributions and Future Research Directions . . . . .	235
8.2.1	Four Challenges of Cognitive Personal Informatics Systems . . .	235
8.2.2	Challenges and Design Considerations for the Design & Use of PI Systems . . . . .	237
8.2.3	Accuracy and Trust of Sensing Technologies . . . . .	240
8.3	Limitations . . . . .	242
8.3.1	Sampling & Selection Bias . . . . .	242
8.3.2	Novelty Effect . . . . .	242
8.4	Conclusion . . . . .	243
	<b>Bibliography</b>	<b>244</b>
	<b>A Fitbit For the Mind?</b>	<b>267</b>
	<b>B Study 1</b>	<b>274</b>
B.1	Participant Information and Consent Form . . . . .	275
B.2	Interview Schedule . . . . .	280
	<b>C Study 2</b>	<b>283</b>
C.1	Screening Questionnaire . . . . .	284
C.2	Participant Information and Consent Form . . . . .	288
C.3	Interview Schedule . . . . .	293
	<b>D Study 3</b>	<b>297</b>
D.1	Participant Information Sheet - Research Summary . . . . .	298
D.2	Participant Information Sheet . . . . .	302



D.3	Demographic Questionnaire . . . . .	314
D.4	Interview Schedule . . . . .	318
<b>E</b>	<b>Study 4</b>	<b>324</b>
E.1	Information Sheet . . . . .	325
E.2	Consent Form . . . . .	328
E.3	Electronic Survey Questions . . . . .	330
<b>F</b>	<b>Data Management Plans</b>	<b>335</b>
F.1	General . . . . .	336
F.2	Study III . . . . .	341
<b>G</b>	<b>Miscellaneous</b>	<b>346</b>
G.1	13-point Ethics Checklist . . . . .	347

# List of Figures

2-1	The stage-based model of Personal Informatics . . . . .	33
2-2	The lived informatics model of Personal Informatics . . . . .	35
2-3	Classification of wearable devices . . . . .	42
2-4	Wearables market share . . . . .	43
2-5	Fitbit Versa . . . . .	53
2-6	Spire Pebble . . . . .	53
2-7	NeuroSky MindWave headset . . . . .	54
2-8	Recording sites for electrophysical signals . . . . .	56
2-9	Clinical EEG recording . . . . .	57
2-10	The basic control and design of a BCI system . . . . .	57
2-11	Medical-grade BCI Speller system . . . . .	58
2-12	Consumer-grade BCI headsets . . . . .	59
3-1	The Cognition Tracker Android application, version 1 . . . . .	70
4-1	Session Recording . . . . .	95
4-2	Session Overview - EEG Metric Summaries . . . . .	96
4-3	Session Overview - HR, BR, Mood and NASA TLX Summary . . . . .	97
5-1	Suggested diagnosis algorithm, showing differences between different fatigue presentation . . . . .	125
5-2	Cognizance Tracker Main Menu . . . . .	127
5-3	Steps of New Recording . . . . .	129
5-4	Session Overview . . . . .	130
5-5	Query Area & Insight Feed . . . . .	131
5-6	Fitbit dashboard showing one-day forward and back data navigation . .	150
5-7	Fitbit User Data Export Page . . . . .	151
6-1	Exist's Correlation Feature . . . . .	158

6-2	Insight Feed - Feed & Insight Details . . . . .	160
6-3	Example Insight Cards . . . . .	161
6-4	User Insight Rating & Insight Variable Information . . . . .	162
6-5	Query Area - Selection & Result . . . . .	167
6-6	Insight Feed Initiation and Usage Over Time . . . . .	172
6-7	Insight Feed & Detail Usage By Participant . . . . .	173
6-8	Insight Feed Usage By Participant . . . . .	174
6-9	Insight Detail Usage By Participant . . . . .	175
6-10	Query Area Usage By Participant . . . . .	182
6-11	Recording vs. Overview vs. Insight Feed vs. Query Area Session . . . . .	185
7-1	Study 4 - Experimental Setup . . . . .	200
7-2	Heart Rate - Participant Expected Trend Graph . . . . .	202
7-3	Four visualisation conditions . . . . .	203
7-4	Data Manipulation Example Graphs - Heart Rate . . . . .	205
7-5	Data Manipulation Example Graphs - Attention . . . . .	206
7-6	Data Manipulation Example Graphs - Meditation . . . . .	207
7-7	Previous BCI and PI experience . . . . .	210
7-8	Confidence in Accuracy of Data Visualisation - Heart Rate . . . . .	212
7-9	Confidence in Accuracy of Data Visualisation - Attention . . . . .	213
7-10	Confidence in Accuracy of Data Visualisation - Meditation . . . . .	214
7-11	Confidence Device is Trustworthy - Fitbit . . . . .	214
7-12	Confidence Device is Trustworthy - NeuroSky . . . . .	215
7-13	Metric Data Visualisation - Closeness to Prior Expectation . . . . .	216
7-14	How confident are you that this device will accurately reflect your level of... ? . . . . .	216
7-15	How confident are you that this device will reflect your...? . . . . .	220
7-16	How confident are you that the device used to collect this data is trust- worthy? . . . . .	220
E-1	Page 1 . . . . .	330
E-2	Page 2 . . . . .	331
E-3	Page 3 . . . . .	332
E-4	Page 4 . . . . .	333
E-5	Page 5 . . . . .	334

# List of Tables

1.1	Thesis Study and RQ Overview . . . . .	25
3.1	Study 1 - Participant Demographics . . . . .	71
3.2	Potential categories of use for CPI systems . . . . .	73
3.3	Suggested metrics/states for tracking using CPI . . . . .	77
3.4	Participant suggested additional data types for combining with EEG. . . . .	80
3.5	Barriers to use and concerns about CPI systems . . . . .	81
4.1	NeuroSky Metric Descriptions . . . . .	93
4.2	NeuroSky Metric Visualisation Types . . . . .	94
4.3	Study 2 - Participant Demographics . . . . .	98
4.4	Summary of Themes . . . . .	102
4.5	Variety of Recording Contexts . . . . .	106
5.1	Study 3 - Participant Demographics . . . . .	132
5.2	Recording Sessions Per Participant . . . . .	133
5.3	Summary of Themes . . . . .	134
5.4	Variety of Recording Contexts . . . . .	138
6.1	Total and Sent Correlation Insights by Participant . . . . .	163
6.2	Caption . . . . .	163
6.3	Total and Sent Trend Insights by Participant . . . . .	164
6.4	Total Extremes Insights by Participant . . . . .	165
6.5	Total Variability Insights by Participant . . . . .	166
6.6	Study 3 - Participant Demographics . . . . .	168
6.7	Feature Usage - Interaction Count . . . . .	169
6.8	Feature Usage - Interaction Length . . . . .	169
6.9	Insight Views By Insight Type and Participant . . . . .	178
6.10	Query Result Views By Query Type and Participant . . . . .	181

6.11	Mean Time from Recording to Sense-making Session . . . . .	186
7.1	Study 4 - Participant Demographics . . . . .	211
7.2	Pearson's Correlations . . . . .	217
7.3	Open-ended Responses - Confidence Device Will Reflect Heart Rate . .	221
7.4	Open-ended Responses - Confidence Device Will Reflect Attention . . .	224
7.5	Open-ended Responses - Confidence Device Will Reflect Meditation . .	226
7.6	Open-ended Responses - Fitbit Device is Trustworthy . . . . .	227
7.7	Open-ended Responses - NeuroSky Device is Trustworthy . . . . .	228

# List of Abbreviations

API	Application Programming Interface
BCI	Brain Computer Interface
BCQ	Brain Conditioning Quantification
CF	Chronic Fatigue
CPI	Cognitive Personal Informatics
ECoG	Electrocorticography
EEG	Electroencephalogram/Electroencephalography
ICF	Idiopathic Chronic Fatigue
iEEG	Intracranial Electroencephalography
IoT	Internet of Things
JSON	JavaScript Object Notation
ME/CFS	Myalgic Encephalomyelitis/Chronic Fatigue Syndrome
MS	Multiple Sclerosis
PI	Personal Informatics
QS	Quantified-Self
SDK	Software Development Kit

## Chapter 1

# Introduction

## 1.1 Personal Informatics

**Personal** *‘Of, relating to, concerning, or affecting a person as a private individual (rather than as a member of a group or the public, or in a public or professional capacity); individual, private; one’s own.’* (Oxford English Dictionary, n.d.-b)

**Informatics** *‘The branch of study that deals with the structure, properties, and communication of information and with means of storing or processing information.’* (Oxford English Dictionary, n.d.-a)

**Personal Informatics Systems** *“[Systems] that help people collect personally relevant information for the purpose of self-reflection and gaining self-knowledge... Effective personal informatics systems help users collect the necessary personal information for insightful reflection.”* (I. Li, Dey, & Forlizzi, 2010, p. 558)

Personal Informatics (PI), sometimes referred to as the ‘Quantified Self’ (QS) (Lee, 2013; Rapp & Cena, 2014) or life-logging (O’Hara, Tuffield, & Shadbolt, 2008), is the process by which people analyse and interpret personal data, which they have collected using digital or paper-based methods (e.g. journals, mood logs, food diaries, fitness trackers), some of which may include the use of external and/or wearable sensors (e.g. smartwatches, step counters, heart rate monitors). QS is described by the Quantified Self (2019) as a movement for “self-knowledge through self-tracking” and describes participation in the QS as “finding personal meaning in your personal data”. Users of personal informatics systems have been shown to be interested in capturing data about themselves for many reasons, including to improve health and other aspects of life, and to find new life experiences (Choe, Lee, Lee, Pratt, & Kientz, 2014), objectives which are aligned with the motivations of the QS movement.

Existing personal informatics research has sought to better understand the contexts within which people use available technologies to collect, analyse and act on data about themselves. For example, in health contexts, research has explored how people use self-tracking and personal informatics methods to come to terms with mental health issues and manage bipolar disorder (Matthews, Murnane, & Snyder, 2017), manage and cope with chronic pain (Felipe, Singh, Bradley, Williams, & Bianchi-Berthouze, 2015), manage symptoms related to multiple sclerosis (Ayobi, Marshall, Cox, & Chen, 2017), and how other chronic illnesses such as ME/CFS may benefit from self-tracking (Davies, Jones, & Kelly, 2019). Self-tracking within the context of health covers a large variety of objectives for collecting and analysing data and spans a broad range of health conditions. Lupton (2020) illustrates previous research efforts that highlight the value of



personal informatics for users seeking to improve physical fitness, moderate or manage healthy eating habits, monitor sexual and reproductive health, and managing mental health and stress.

People who participate in self-tracking or using personal informatics systems may not necessarily identify as members of the niche QS movement. The adoption of self-tracking and personal informatics applications is now becoming part of mainstream culture and a common activity across a broad range of demographics. For example, a focus group study with girls, aged 14-17, by Depper and Howe (2017) found that children believed that health and fitness applications could be beneficial in improving individual focus on health outside of the school environment. Pink, Sumartojo, Lupton, and Heyes La Bond (2017) found that cyclists who used apps to track their cycling habits built up exercise routines based around the data that they had tracked. In relation to reproductive health Epstein et al. (2017) found that apps that enabled tracking of menstrual cycles enabled users to “understand their bodies and mental states, to have materials prepared for their period, to predict ovulation, and/or to describe their menstrual cycle to their doctor” (Epstein et al., 2017).

Didžiokaitė, Saukko, and Greiffenhagen’s (2018) research sought to better understand how ‘ordinary people’, rather than people who may consider themselves part of the quantified self movement, used calorie tracking in their lives. From their interviews with participants they found that many people benefited from the ability to use tracked data “to re-evaluate the foods based on their caloric and nutritional value, but also became generally more aware of what they were eating” (Didžiokaitė et al., 2018). Work by Macleod, Tang, and Carpendale (2013) reported how people living with chronic illnesses can benefit from tracking data about their condition. The process of engaging with PI applications to collect a variety of data (e.g. episodes of illness, triggers, medications, current health status, and history) enabled them “to understand their conditions and regain control over their own lives, which gave them a means to communicate more effectively with their [healthcare] provider and receive better treatment” (Macleod et al., 2013).

PI systems are used by a diverse group of people to track varied data that may help them in better understanding their own lives. PI system research has a large focus in the domains of health and well-being, with people exploring how data can help them to better understand themselves, their conditions, and provide insights which they may then be able to use to change behaviours or make adjustments to their daily lives. The set of people that are keen and able to track data about individual aspects of their lives is therefore expanding to include larger swathes of the population and is not

just limited to those who are “technophiles”, i.e. those who are enthusiastic about new technologies. Rather, they include the ‘ordinary’ people who, as described by Pink et al. (2017), may be tracking the ‘mundane’. However, the increasing breadth and variety of data that different people are tracking means that PI systems should sufficiently support people, with varying levels of knowledge, in their attempts to make sense of that data. People’s levels of domain knowledge about the data they are collecting will vary as will people’s knowledge of interpreting data using appropriate statistical methods. However, as much as there is variety in types of things that may be tracked in relation to one’s health or well-being, there is also variety to the other areas of people’s lives in which PI systems may be useful.

Beyond the health context, recent personal informatics research has explored how PI systems can be used to improve productivity. TimeAware was developed to understand the ways in which an application capable of tracking personal productivity data can encourage people to increase their productivity (D. J. Kim, Lee, Rho, & Lim, 2016). The application enables reviewing of time spent in different levels of productivity by building upon data available from RescueTime<sup>1</sup> - a tool that can be used to track time spent in different applications and websites. TimeAware allowed the captured productivity data to then be visualised for participants in the study, through widget and dashboard views. D. J. Kim et al. (2016) found that the tool was successful in promoting productivity when results were negatively framed i.e. drawing attention to the amount of time spent in ‘distracting activities’. A. N. Meyer, Murphy, Zimmermann, and Fritz (2017) also sought to better understand how work productivity might be improved by conducting design evaluations and a field study with software developers. The participants supplied self-reported productivity levels via a pop-up at intervals throughout the day. Insights were then generated and presented to participants after data collection. The authors conclude that not only did providing self-reported productivity levels generate a degree of self-awareness of the participants’ productivity, but participants were surprised by some of the insights that were generated, such as the amount of time spent working collaboratively with others.

The data that one might initially consider tracking with PI systems is not always the obvious signal (e.g. recording “productive time” for reflection on productivity) as in the previous study. Sometimes, tracking external factors or surrogate markers can reveal previously hidden relationships between variables and provide actionable insights. For example, Rivera-Pelayo, Fessler, Müller, and Pammer (2017) assessed the potential for *mood tracking* applications to improve employee *performance* in call centres and found

---

<sup>1</sup><https://www.rescuetime.com>

that there was a benefit to be gained in terms of improved key performance indicators. They suggest that mood tracking can be beneficial if the tracking is sufficiently well integrated within the work context (e.g. used where employees are already monitored, individuals are involved in the decision to provide mood information, and the utilisation of this data suits the management style). With research efforts to develop sensors that are capable of detecting a wearer’s mood, the collection of such data could become automated in the future. A meta-analysis conducted by Malhi et al. (2017) highlights the state-of-the-art and limitations of mood tracking, and suggests that there is a potential for heart rate variability and skin conductance measurements to be used as bio-markers of mood. Despite this being a nascent research area, we are already beginning to see commercially available technologies that promise the ability to detect mood such as Feel<sup>2</sup> and Upmood<sup>3</sup>. Both of the aforementioned devices are wrist-worn wearable sensors that claim to be able to track a variety of emotions.

Further examples of the proliferation of PI systems can be found in the context of finances, where self-tracking applications enable people to track income and expenditure (Epstein et al., 2016). The emergence of smart home sensing devices enables reflection and behaviour change in relation to consumption of utilities, such as water, gas and electricity (Yang, Newman, & Forlizzi, 2014). Gulotta, Forlizzi, Yang, and Newman (2016) highlights the inclusion of “a core PI feature within the Nest<sup>4</sup> smart-thermostat system, recording data about users’ interactions with the system and then using that information to try and make energy-saving decisions on their behalf” (Gulotta et al., 2016).

As can be seen in the above examples, the breadth of PI application areas extends well beyond health and covers a variety of different aspects from one’s life, including broader aspects of well-being, emotions, finances, and productivity, to name a few. The expansion of PI application areas is in part due to ongoing research efforts that seek to understand what ways recording data may be useful, but can also be ascribed to advances in automated sensing and detection.

The proliferation of technologies that can ‘sense’ and collect data that may be burdensome or infeasible to track manually, such as continuous mood logs or interactions with a device, provides opportunities for users to track increasing volumes of data without expending additional effort in data collection.

Significant challenges in the field of personal informatics include: addressing sense-

---

<sup>2</sup><https://www.myfeel.co>

<sup>3</sup><https://upmood.com/>

<sup>4</sup>[https://store.google.com/gb/category/connected\\_home](https://store.google.com/gb/category/connected_home)

making obstacles that people encounter in relation to the data that is collected by self-trackers (S. Jones & Kelly, 2016); developing applications for those new to self-tracking, who may be more likely than experienced trackers to abandon tracking devices, struggle to find useful information in the data, or find the tracking too burdensome (Rapp et al., 2018); as well as exploring how to best collect user-generated data during in-the-wild studies (Vaizman, Ellis, Lanckriet, & Weibel, 2018).

## 1.2 Motivation for this Research

The availability of data collected by and about a person is increasing. This is in part due to the availability, miniaturisation and lowering cost of sensing technologies. Previously, to track step counts required the use of a pedometer (e.g., Lin, Mamykina, Lindtner, Delajoux, & Strub, 2006) or a standalone ‘activity tracker’, but now this ability is built into most modern smartphones or smartwatches, such as Apple Watch and Fitbit. Technologies that initially work as stand-alone products are becoming sufficiently small to be integrated within multi-purpose devices, thus reducing cost to consumers by not having to own multiple devices, and cost to manufacturers in terms of the amount of materials used.

As an example of the increasing availability of sensing technologies, prior work by Zheng et al. (2014), highlighted a variety of potential ways in which heart rate could be measured, including finger-worn rings, in-ear and on-ear based sensors, sensors built into glasses, glove- and hat-based sensors, and smartphone-based sensors. Several of these technologies have since been developed into commercially available standalone items such as the Oura ring<sup>5</sup>, or provided in addition to other technologies such as the Jabra Elite in-ear headphones<sup>6</sup> and Solos smart glasses<sup>7</sup>. Research also continues to investigate other ways in which sensing technologies can be integrated into everyday objects, including within research in emerging fields such as smart-textiles (Niknejad, Ismail, Mardani, Liao, & Ghani, 2020).

The purpose of the research presented in this thesis is to develop an understanding of the issues that users of personal informatics systems encounter when tracking data from newly emerging sensing technologies that are *low cost*, *unobtrusive*, may be *wearable*, and the technologies produce data that is *personal*, and *readily available*. Furthermore, the data captured by these technologies produce large volumes of a variety of data that may be *multifaceted*, and users may have little-to-no previous experience with the

---

<sup>5</sup><https://ouraring.com>

<sup>6</sup><https://www.jabra.co.uk/sports-headphones/jabra-elite-sport>

<sup>7</sup><https://solos-wearables.com>

data produced. The *trustworthiness* of the captured data may not be immediately be apparent, and the mechanisms and processes through which data values have been ascertained or calculated by the technologies may not be *transparent*. A more detailed definition of the characteristics of the technologies that are considered within the scope of this research is provided in Chapter 2, Section 2.3.

As can be seen in the previous examples, there are a plethora of ways in which people can now track and review data they have collected about themselves. This proliferation is not only limited to self-motivated tracking, but may also might be encouraged by others, such as companies or governments. This is described by Lupton (2017) as ‘communal self-tracking’, rather than ‘private self-tracking’. Examples of this include life and health insurance companies basing premiums on data recorded by fitness tracking data (McFall, 2019), and schools mandating the use of EEG sensor tracking applications to determine children’s levels of attention during classes, as reported in news outlets such as the *Wall Street Journal* and *People.cn* in 2019.

Given these examples, it is possible that in the future a large portion of society will be making use of technologies that allow for the collection of, and reflection on, data that is related to them. It is also likely that they may want to combine and reflect on multiple sources data in combination, rather than reviewing siloed data in isolation (Videnova, 2016; Epstein, Cordeiro, Bales, Fogarty, & Munson, 2014).

At present, little is known about how people use and attempt to make sense of data from emerging personal informatics technologies that are yet to become pervasive in society. A large portion of research within PI seeks to understand the ways in which tracking data may help people to better understand specific aspects about themselves. This is usually done by researching how people are using currently available sensing technologies or whether currently available sensing technologies can be re-used to provide benefit in a new context. However, little research focuses on how people use emerging sensing technologies to learn more about themselves. In particular, what sense-making do people attempt to perform when given a new sensing technology. We need to further understand how people might attempt to make sense of data from these technologies when the benefits of the sensing technology itself is not being studied. This need is due to the increasing availability of sensor technologies and apps that allow users to integrate the data from them. In particular little attention has been paid to the ways in which mainstream technology users attempt to make sense of the data these technologies provide, except for some recent studies seeking to address the ‘mundane’ tracking (Pink et al., 2017; Didžiokaitė et al., 2018) as described previously.

Another factor that will likely influence the future of PI systems is the nature by which multiple sensing technologies may be used by a single person, or at least they may have the ability to track multiple streams of data from one piece of technology. It is still under-explored how combining data from multiple sensing technologies may be investigated and interpreted differently by people and require different approaches to uncover insights rather than being provided with a specific sensing technology designed to track one aspect of one's life in isolation. We need a deeper understanding of how people go about trying to make sense of these data in combination with each other rather than in isolation, if multiple sensing technologies are being used to provide insights.

To answer these questions we have conducted research using novel sensing technologies, with participants who have little-to-no previous experience with self-tracking. The research studies seek to understand both context-specific and context-free use cases of how people go about tracking and then attempt to make sense of the collected data within these contexts. As the reasons people decide to track vary, the context-free use case will have no specific requirements for inclusion. In the context-specific use case participants will be required to have been diagnosed with a specific form of fatigue for which they feel they may gain benefit from tracking data. The studies also explore whether there are specific application features which may encourage people to investigate their data further, or whether there are features which may provide easier access to insights related to the collected data. The research described in this thesis takes place as both lab-based and in-the-wild studies. Lab-based studies enable us to conduct research in a more controlled environment to answer specific questions about the technologies being studied. The controlled nature of lab-based studies allows us to exclude external factors that may be encountered by participants outside of a lab setting. In-the-wild studies on the other hand, enable us to gather results which are situated in the everyday way people use the technologies and indicates how people might use the technologies in the real world. A majority of the research results presented in this thesis are qualitative results, based on the analysis of responses from semi-structured interviews. This method will allow for participants to provide more open-ended responses, and ensure that participants' responses can be fully understood. This will allow us to capture feedback that may be outside of our initial hypotheses, that may not be captured in structured survey responses or quantitative methods of data collection.

### 1.3 Thesis Structure & Contributions

This thesis provides the following contributions to the developing field of personal informatics and in particular the emerging field of novel sensing technologies within the context of PI systems. It provides an understanding of how people analyse and use data recording from emerging sensing technologies by way of lab-based technology probes and longitudinal in-the-wild studies of fully-functional personal informatics applications. Qualitative and quantitative research presented here provides an understanding of the potential benefits, pitfalls and personal requirements of the types of people that might be inclined to make use of sensor data to better understand themselves. Additionally this research provides the results of novel means of allowing people to analyse their data by way of predefined queries and statistically generated insights. An overview of how each chapter and study relate to each other can be found in Table 1.1.

Chapter 2 provides background information on personal informatics, current sensing technologies used within PI and brain-computer interfaces, and their applicability as an emerging sensor technology within the context of this research.

Chapter 3 discusses an exploratory user study that aimed to understand the possibilities and challenges presented by using a BCI as a sensor within the context of a personal informatics application capable of recording and viewing BCI data.

In Chapter 3 we gain a better understanding of the ways in which people might expect to use a novel sensing technology. The exploratory nature of the chapter’s research provides groundwork for the development of a feature-rich application that contains features and metrics that people appeared to show an interest in or willingness to explore further, in relation to PI. The chapter contributes challenges and design considerations for the further exploration of novel sensing technologies, in particular in relation to BCI technologies. We investigate these challenges further using the developed feature-rich mobile application throughout the research described in Chapters 4, 5, and 6.

Chapters 4 and 5 build upon the challenges, presented in Chapter 3, by way of an in-the-wild study of a PI application capable of recording and visualising EEG data as metrics as well as pulling in data from secondary data sources: heart rate data from Fitbit and breathing rate data from Spire. The two studies address two different groups of potential PI system users. The study discussed in Chapter 5 builds upon some of the work of the previous studies but also implemented specific design considerations of the target population, those living with ME/CFS, CF, and ICF.

In Chapter 4, we provide a further understanding of how people go about recording,

Study 1 - Chapter 3	
Aim	An exploratory study to elicit views on how and why cognitive personal informatics system might be used, and to capture expectations for features that future systems may include.
RQs	RQ1.1 What are the design opportunities, challenges, and technical, social and ethical implications for the “near-future” technology of ‘Cognitive Personal Informatics’?
Methods	An exploratory user study using a PI technology probe app, participant interviews and thematic analysis.
Challenges	C1: Addressing the indirection between meaningful psychological states and brain activity data
	C2: Supporting diverse tracking styles
	C3: Encouraging exploratory & enquiring approaches
	C4: Overcoming misconceptions & lack of understanding about the brain
Study 2 - Chapter 4	
Addresses	C1: Addressing the indirection between meaningful psychological states and brain activity data
Methods	An in-the-wild user study using a PI app created using suggestions from Study 1, participant interviews and thematic analysis.
RQs	RQ2.1 How do users record, collect, analyse and make sense of data from ‘novel sensor technologies’ using a multifaceted PI system?
	RQ2.2 What are the implications of providing people with metrics which have not yet been clinically validated?
Challenges	C5: Trust in PI (Explored in Study 3a and 4)
	C6: Simplification of Insight Extraction (Explored in Study 3b)
Study 3a - Chapter 5	
Addresses	C2: Supporting diverse tracking styles, C5: Trust in PI
Methods	An in-the-wild user study using a PI app with modification suggested from Study 1 and modifications for the participants concerned, participant interviews and thematic analysis.
RQs	RQ3.1 How do people with ME/CFS perceive benefits and limitations of collecting data from novel sensing technologies?
	RQ3.2 How do participants attempt to make sense of data from a range of novel sensors?
	RQ3.3 How do people determine if a device is trustworthy?
Study 3b - Chapter 6	
Addresses	C3: Encouraging exploratory & enquiring approaches, C6: Simplification of Insight Extraction
Methods	As with Study 3a
RQs	RQ3.4 Does an ‘Insight Feed’ of pre-computed insights, encourage further exploration and sense-making of data?
	RQ3.5 Does providing a predefined set of possible queries, as in the Query Area, enable people to better explore and make sense of their recorded data?
Study 4 - Chapter 7	
Addresses	C5: Trust in PI, Further explores trust in PI based on the recurring theme encountered in prior studies.
Methods	Randomised lab-based study using MindWave and Fitbit wearables. Quantitative and qualitative analysis of participant responses.
Hypotheses	H1 There will be a significant increase in user trust in data when the data more closely reflects users’ prior predictions about data values.
	H2 There will be a significant difference in user trust in a personal informatics device and the data it produces according to the precision at which it is presented within a visualisation. Therefore, rounded values will be less trusted than precise values.
	H3 Trust in the sensor data will differ based on the level of knowledge the participant has about how a particular sensing technology is measuring and capturing data.

Table 1.1: Thesis Study and RQ Overview



analysing and attempting to make sense of data from novel sensor PI systems in-the-wild. The need to provide novel sensing technologies that are discreet and unobtrusive to the user is highlighted. During this study we find that users of novel sensing technologies in the context of PI may benefit from the simplification of insight extraction from their data and may be aided with knowledge transfer from domain experts to them. We also highlight the need for users to have means to assess the accuracy of their devices and an initial understanding of trust in relation to novel sensing technologies is discussed.

In Chapter 5, we further explore the ways in which people living with prolonged fatigue, chronic fatigue (CF), idiopathic chronic fatigue (ICF) or myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS) might attempt to use a novel sensor PI system to better understand themselves and their condition. We find that people in this study have unique requirements that take their illness into account and consider the time-frames for which they wish to track data. These findings should be considered when developing PI tools for those with chronic conditions or specific reasons to track for health reasons. Further design considerations, supporting variety in data presentations and providing information about how a sensor’s data is formed, are presented that may provide benefit to those tracking PI data.

To shed some more light on the challenge of ‘Encouraging exploratory and enquiring approaches’ described in Chapter 3, in Chapter 6, we discuss two novel ways for users to discover insights about their recorded data. The first approach uses automatically generated statistical insights about a user’s recorded data in the style of a ‘News Feed’. The second way of uncovering insights was via a set of predefined statistical analyses situated within a ‘Query Area’. This chapter provides the qualitative results of this research which was undertaken with the same group of participants from Chapter 5. In Chapter 6 we contribute design consideration for the development of PI system features which attempt to bring information of interest to the user.

In Chapter 7 we present the results of a lab-based study that sought to better understand how peoples’ formation of trust in a sensing technology’s data may be influenced based on how the collected data is visualised. This chapter contributes a further understanding of the role data visualisations of recorded data might play in affecting peoples’ trust in PI sensing technologies; as we found during the previous studies that trust may be influenced by inherent trust in technology and/or trust in the sensor manufacturer or vendor. Participants’ responses from these studies also suggest that their trust in the technology may be questioned when the values do not meet users expectations and that perceived accuracy of data may also be another factor. Our quantitative findings from

this study suggest that trust is likely to be set prior to their analysis of representations of their data, such as data visualisations.

In Chapter 8, we conclude by revisiting this thesis’s goals and contributions based on the previous chapters. In addition we present limitations of our research and directions for future work that may be undertaken.

## Chapter 2

# Literature Review

## 2.1 Chapter Overview

This chapter presents background research related to Personal Informatics (PI) research areas, systems and sensing technologies. We highlight what properties or characteristics we consider ‘novel’ sensing technologies to possess and provide a discussion regarding our choice of a BCI as a sensor that exemplifies these characteristics. Based on this sensor choice we provided further background information regarding how this technology works, as well as information about other sensing technologies used in this research, namely heart rate and breathing rate sensors. Additionally, this chapter explores current research related to the topic of trust within the context of PI.

Section 2.2 provides additional background concerning PI, in particular PI systems and models of PI, expanding on PI as described previously in Chapter 1, Section 1.1, as well as the current state-of-the-art research within PI. Sections 2.3 and 2.4 discuss the role sensors availability plays in self-tracking, the different forms they take, and their availability for commercial and research purposes. Section 2.5 explains our choice to use a commercial brain-computer interface device as a technology probe. Therefore, Section 2.6 provides background information relating to brain-computer interfaces, including current medical uses of BCIs. This is followed by information provided in Section 2.7 about BCI research related to self-tracking and behaviour change in fields becoming known as neurotechnology and neuroergonomics. Finally in Section 2.8 we describe current information about the role trust plays in PI.

## 2.2 Personal Informatics

As discussed in the introductory chapter, Personal Informatics may be considered as the branch of study related to using information about a person, by that person, for self-reflection and gaining self-knowledge. Personal Informatics in its current state has overlapping research interests with life-logging and self-tracking. A large portion of PI research is focused on health and medicine, however, as described previously that is not necessarily the entirety of PI research with PI tools being used in areas ranging from health, well-being, finance and productivity. In addition, the range of things that people can track is increasing with the proliferation and increasing availability of technologies that enable people to track and log facets of their lives.

### 2.2.1 Personal Informatics Research

To further understand the current state of PI research, we first look to initial work related to behaviour change, in particular research utilising automated sensing. Before the rise of the Personal Informatics movement systems which used the increasing ubiquity of mobile technologies to capture data about people were being researched to determine their applicability in the domain of behaviour change, in particular in the area of health. These tools include using mobile phone-based applications to manually capture data, augmenting computer-based technologies using pedometers, or using pervasive computing data sources, such as mobile phone signal, as surrogates for physical activity data.

An example of this is ‘Fish’n’Steps’ developed by Lin et al. (2006), an interactive computer game used in combination with a pedometer, who sought to determine its feasibility in encouraging physical activity via data visualisations contained within a ‘social video game’. Within the game the participants’ daily step count was linked to the growth of a virtual fish in a fish tank, some of the other players’ fish were also included in an individual’s fish tank which encouraged an atmosphere of cooperation and competition. The study found that this type of interaction encouraged participants to perform more physical activity, but participants found that the pedometer and the requirement of manually uploading the data at a public kiosk inconvenient (Lin et al., 2006).

Tsai et al. (2007) developed a mobile application (PmeB) within which users could record their caloric intake, with the self-improvement goal being increased weight loss. The mobile application allowed users to record caloric intake by entering foods and provided users with a visualisation of their remaining calories. An additional pedometer allowed participants to enter their step count, which was converted to calories burnt. The researchers evaluated the use of prompts to encourage users to make entries and found increased compliance when compared to a control group using pen and paper to track their caloric intake.

Maitland et al. (2006) developed Shakra, a mobile application designed to increase awareness of daily physical activity. The Shakra mobile application used changes in the users’ mobile phone signal to estimate their activity level, i.e. walking versus stationary. The application integrated visualisations that allowed users to compare their activities by day and against other users of the application. They found that this application increased users’ awareness of their activity levels and provided encouragement to be more active.

These examples may be considered as little steps in the direction which we now find ourselves where PI systems are used in a variety of different situations including trying to increase physical activity, tracking food for dieting, tracking income and expenditure to manage their personal finances, and tracking computer and phone usage to increase productivity. Users are now aided by the ubiquity of smartphones, smartwatches, and wearable activity trackers. Although research continues to look at ways which new technologies might aide people in behaviour change, current PI research can be considered under the following objectives: understanding tracking practices, developing recommendations for design features, and providing visualisation techniques, which will be discussed in the following sections.

### **2.2.2 Understanding Tracking Practices**

This area of research seeks to understand how people use currently available tools, why they use them, and what they hope to gain from self-tracking. The objective of this research is to provide insights into the ways in which people are using currently available methods of tracking to better understand themselves.

In conducting an interview study of people with chronic illnesses Macleod et al. (2013) suggests that by focusing on curiosity rather than behaviour change, as well as the role PI tools play in peoples' lives, it will be possible to develop tools that support people in treating and managing chronic illness. People with illnesses such as asthma, depression, diabetes, epilepsy, fibromyalgia, gastroesophageal reflux disease, showed similar interest in better understanding themselves to help manage their illnesses as the PI enthusiasts of I. Li et al. (2010) and I. Li, Dey, and Forlizzi (2011) did in tracking aspects of their lives.

Choe et al. (2014) attempted to understand what motivates “extreme users” of PI systems to track data. By conducting qualitative and quantitative analyses of users of PI systems they gained insights as to why these quantified-selves track certain aspects of their lives, what tools they use to track it, and what they learned from tracking those aspects. They also outline difficulties that users often have in self-tracking, such as: tracking too many things, not tracking triggers and context, and not performing their analysis with sufficient scientific rigour. They provide suggestions for future research such as: enabling early feedback to users to identify what to track and maximising the benefits of manual tracking to increase self-reflection, as it was felt too much automation may reduce the awareness and self-reflection.

Further work by Choe, Lee, and Schraefel (2015) characterises the types of visualisa-

tions that members of the quantified self movement use and what insights they hope to gain. These are broken down into eight visualisation types: Detail, Self-Reflection, Trend, Comparison, Correlation, Data Summary, Distribution, and Outlier. These are then further categorised into subtypes, such as identifying extremes, values, and references within the category of *Detail*. Self-reflection can be broken down into subtypes of external context, contradiction, prediction, and confirmation. Choe, Lee, and Schraefel (2015) also provides an overview of the most frequently used visualisations, finding that line charts, bar charts and scatter plots were the three most common.

Ayobi et al. (2017) specifically sought to understand the tracking habits of those with multiple sclerosis (MS). They found that those with MS used a variety of tools to track different aspects of their lives to help them manage their MS. They used tools to track their physical activity to remain active, diaries to enable them to track and manage their MS symptoms, as well as medication tracking. Participants also tracked data in life journals for improved mental health.

These research topics provide relatively fine-grained details about how people use self-tracking technologies with the aim of understanding how specific groups of people may use self-tracking as a means to better understand themselves. This area of research tends to have a specific user group, based on what it is they are trying to achieve and how they achieve this. This uncovers information that others may find beneficial if they were to start doing it, or may provide insights into how to best develop tools in future.

To better understand the way in which people more generally attempt to track data in relation to PI, researchers have attempted to model the processes of self-trackers. The current models that are most widely used to describe this process are discussed in the following section.

### 2.2.3 Personal Informatics Models

Further to understanding the tracking practices of individuals other research seeks to model the behaviours of people engaged in PI. Two models currently tend to be used to discuss the processes of PI users, they are the *Stage-based* and the *Lived-informatics* models. These models express the stages that a person may transition through when they begin and proceed through the process of tracking for self-understanding. The models can be used to guide the effective design and implementation of self-tracking tools by providing a better understanding of the steps a person who will use the tools may take. These models also enable researchers to better describe the stages that

their research ideas attempt to address. The stage-based model was developed to better understand the problems that people encounter when using PI tools, as at the time very little was understood about this. The lived-informatics model extends the stage-based model to account for additional stages that are not considered in the stage-based model by building upon the characterisation of PI as “lived informatics”. An understanding of PI as ‘lived informatics’ provides insights into how PI tools are used depending on a person’s reason, or goal, for self-tracking, as well as to obstacles that may need to be overcome.

### 2.2.3.1 Stage-Based Model

The stage-based model of PI (see Figure 2-1) suggests that users of PI systems follow an iterative five-stage approach consisting of preparation, data collection, data integration, reflection, and action. The authors developed this model by surveying and interviewing participants who “collect and reflect on personal information” (I. Li et al., 2010). The initial preparation stage is where, based on a specific motivation, a user decides what data they would like to collect and how they are going to collect it, often making use of wearable or smartphone tracking technologies. The subsequent collection stage is where users collect and log the required data. The integration stage then refers to the process of preparing the data in such a way that it may be reflected upon. Users can then reflect on their collected data; the reflection stage may be short- or long-term. The action stage then sees users deciding what actions to take based on their gained self-knowledge. The researchers highlight these stages as barriers to progression and should be considered as a way of improving ‘the diagnosis, assessment, and prediction of problems in personal informatics systems’ (I. Li et al., 2010).

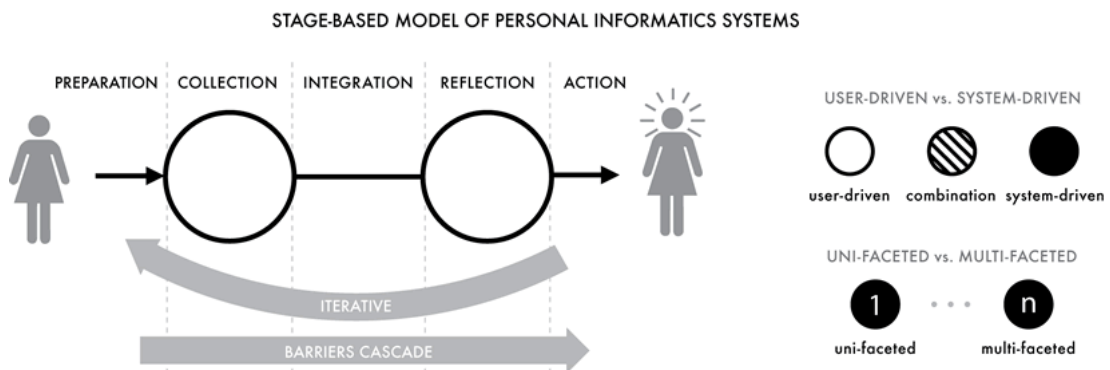


Figure 2-1: The stage-based model of Personal Informatics (I. Li et al., 2010)

In addition to modelling the stages people move through when engaging in PI, later work by I. Li et al. (2011) describes six kinds of questions that people asked about their



personal information. The questions are defined by I. Li et al. as *Status*, *History*, *Goals*, *Discrepancies*, *Context*, and *Factors*. Users examining their *Status* are seeking to know the current value of what they are tracking, so they see how they are in the present moment. *History* allows users to answer questions by seeing data over the long-term to find trends and patterns. Users try to find what *Goals* are appropriate to pursue based on their current data. *Discrepancies* are found by comparing a user's current status with a specific goal. *Context* helps to answer questions about what things are happening at or near the same time as the user's current-information seeking. *Factors* provide information about what influences behaviour over a longer period of time. This study also defines two phases of reflection, *Discovery* and *Maintenance*. During the Discovery phase users do not necessarily know what goal they are trying to meet or the factors that influence their behaviour, whilst during the Maintenance phase their data is used to maintain an awareness of their status relative to their goal (I. Li et al., 2011).

### 2.2.3.2 Lived-Informatics Model

The characterisation of PI as 'lived informatics' by Rooksby, Rost, Morrison, and Chalmers (2014), believing Li et al.'s stage-based model to be too technology-centric, found five overlapping styles of use of technologies within the context of PI: *directive tracking*, *documentary tracking*, *diagnostic tracking*, *collecting rewards*, and *fetishized tracking*. Directive tracking implies that users choose to track data with a focus on achieving a particular goal. Documentary tracking is a style of tracking whereby users track personal data to document their activities, rather than change them. Diagnostic tracking is expressed where a user tracks data to discover an association or understand a relationship between two separate facets. Collecting rewards was noted as being when users track data to score points or register achievements. Fetishized tracking is where users track data purely because they have an interest in technology (Rooksby et al., 2014). The categorisation of these styles of tracking allows us to consider how specific technologies might be designed and developed to help people with specific goals in mind, rather than attempting to find a one-size-fits-all approach.

The 'lived informatics' model of tracking, shown in Figure 2-2, extends I. Li et al.'s stage-based model to account for lapsing and resuming in the use of tracking devices and builds upon the characterisation of PI as lived-informatics. This model breaks up the preparation stage of the stage-based model into *Deciding* what to track and *Selecting* the appropriate tool for the task. The stages of Collection, Reflection and Integration, from the stage-based model, are covered under Epstein's heading of *Tracking & Acting*.

At this point users *Lapsing* from capturing of data may occur. However, *Resuming* of data collection can also occur at this point, leading directly back into the stage of Tracking & Acting (Epstein, Ping, Fogarty, & Munson, 2015).

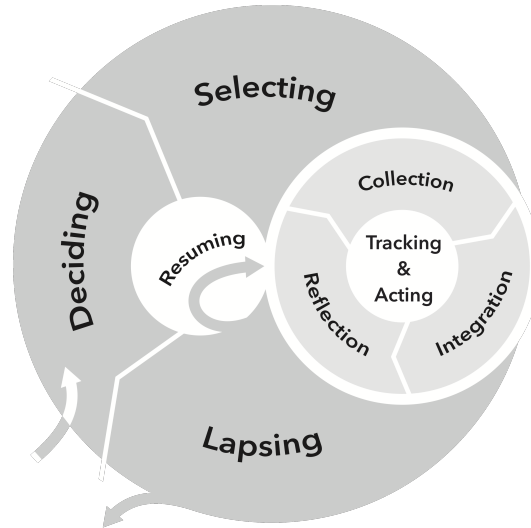


Figure 2-2: The lived informatics model of Personal Informatics (Epstein et al., 2015).

#### 2.2.4 Recommendations for Design Features

Another area of research within the field of PI seeks to provide recommendations for design features of PI tools. To better understand this aspect researchers often perform systematic reviews of currently available tools or qualitative research with users of existing or new tools to understand the implications of features' designs.

Epstein et al. (2017) conducted research to uncover design opportunities within menstrual tracking. To do so they conducted a study which evaluated currently available mobile apps, as well as performing surveys and qualitative interviews with women to understand their tracking practices. This allowed them to suggest implications and recommendations for future tools in this domain, such as: avoiding gendered colours, text and iconography, as well as supporting varied reasons for tracking, and supporting migration between tools.

Ravichandran, Sien, Patel, Kientz, and Pina (2017) undertook a study involving sleep experts, users of sleep sensing devices, and consumer reviews to determine design opportunities in sleep sensing technologies such as displaying data in ranges rather than single-point values, making the sleep sensing algorithm more transparent, focusing on actionable feedback that integrates modifiable behaviours, and personalising feedback.

Eikey and Reddy (2017) provided an understanding about the design of weight loss applications, by conducting surveys, think-aloud exercises, and semi-structured interviews with their participants to determine the uses of, and problems with, weight loss apps. They found that weight loss apps can be used for eating disorder recovery but also may contribute to eating disorders, due to the heavy focus on numbers encouraging obsessive logging.

Kelley, Lee, and Wilcox (2017) conducted focus groups and card sorting exercises with student health professionals and conducted an online survey with 297 students both with and without diagnosed mental health problems. This resulted in “opportunities in leveraging self-tracking for mental wellness, highlighting several design considerations”(Kelley et al., 2017). The student health professionals back up the claims of Eikey and Reddy (2017), concluding that over-monitoring may have a detrimental effect on student well-being. However, they provide suggestions that there is an issue knowing what to track for mental wellness due to lack of correlates, unlike physical activity and fitness. The majority of student respondents tracked behavioural proxies rather than mood directly. An additional consideration is “Designing for Student Life Changes”, noting that for people of other ages or in different situations the implications of their research may not apply.

### **2.2.5 Visualisation Techniques**

A specific feature in PI applications is graphs and charts of a person’s collected data. Therefore researchers also look at methods of visualising PI data such that it can be understood better or may have a greater influence on a users’ self-understanding and -improvement.

Fan, Forlizzi, and Dey (2012b) provides the framework for the visualisation of physical activity as abstract art, Spark. This type of visualisation was used to determine the effectiveness of abstract art as a means of encouraging physical activity in older people (Fan, Forlizzi, & Dey, 2012a). They found that the glance-able nature of the abstract visualisation on a tablet encouraged their participants to be more active. It also encouraged sharing, with one participant showing Spark to their visitors.

StepCity (Walsh & Golbeck, 2014), is a prototype social game developed to encourage users to take more steps. This work builds, in part, upon the Fish’n’Steps research of Lin et al. (2006) mentioned earlier in this chapter. Participants were split into three groups; a control group, a basic social interaction group and a social game group. Participants in the social game group used Fitbit steps as currency to buy buildings

for their virtual town. Participants in the basic social interaction group were just able to see their friends' steps and send messages. They found a causal link for the game encouraging newer users of Fitbits to take more steps compared to a control group who wore the Fitbit as normal.

A study by Epstein et al. (2014) provided participants with visualisations of their data, showing subsets of data with shared features (referred to as 'cuts'), for example, location and physical activity. They sought to understand participants' preferences for presenting these cuts by providing the data in a variety of formats such as graphs, tables, maps, and Sankey diagrams. These cuts were then compared to the visualisation provided by a readily available smartphone app, Moves. Participants' beliefs varied on how factors influenced their behaviour but believed that these cuts could help lead trackers to new discoveries, finding that self-trackers can identify correlations and opportunities for self-improvement by providing them with cuts of data.

#### **2.2.6 Personal Informatics Systems**

The group of tools used by people to collect this self-relevant information tend to be described, generally, as 'Personal Informatics systems'. These systems come in a variety of different types, from paper logs to fully-fledged mobile apps and online dashboards. PI systems usually encompass both the tools that may be used to track the data, as well as tools that may be used in some form to analyse the self-relevant data.

Paper-based methods of recording data are often seen in the form of diaries and logs. This method of logging may be used by people who are non-technical or for recording data that either does not require technology or does not yet have a technological solution suitable for an individual's needs. Similarly, people may use non-technological methods for the analysis of their collected information.

Some well-known health-related PI apps include Fitbit, Google Fit, Apple Health, MyFitnessPal. These applications provide users with the ability to track a variety of data related to health such as heart rate, step count, minutes active. Fitbit and MyFitnessPal are two of the more domain-specific applications. Generally, people use Fitbit for tracking steps, heart rate, and sleep. MyFitnessPal is used to keep track of food intake and thus calories consumed. The data can then be reviewed in the application allowing the user to determine if they need to take more steps, whether they have reached their specific calorie intake goal, or whether they have achieved an optimal number of hours of sleep. The goals or targets are specific to a user, although some may be based on best practices such as achieving 10,000 steps per day or eight

hours of sleep per night, these defaults can be changed and set specifically for the user, as an individual. Google Fit and Apple Health are less specific and provide a single point of access for someone to their data. Google Fit is described as a “health tracking platform” and largely provides APIs that enable developers of other applications to integrate into the Google Fit platform, Apple Health operates similarly. Both Google and Apple have smartwatches which enable tracking of data similar to that which can be obtained from Fitbit but their platform solution potentially provides a one-stop shop for someone’s personal information access regardless of what sensor or device someone chooses.

Gyroscope<sup>1</sup> is another mobile application allows for the tracking of a variety of different data sources such as “places, exercise and even your mood”. The application also enables data from third-parties, such as Fitbit for steps and heart rate data, to be pulled into the app.

Some of the previously mentioned mobile applications come with a web-based tool, however, there is also a type of PI tool that allows for the collation of multiple individual PI data streams and provides additional features because of this. One example includes Exist.io<sup>2</sup> which allows users to connect multiple sources of personal data and then provides an additional feature of being able to determine correlations between disparate data sources.

How people track their data is often dependent on what it is they are trying to track. For example, although it is not impossible for someone to manually log their steps over an entire day, someone is much more likely to use an automated system such as a pedometer, Fitbit or mobile phone to assist them in doing this and therefore must rely on a technological solution to their problem. However, tracking mood, medication, symptoms of an illness may be more easily be dealt with using paper-based methods. If someone requires instant, or immediate, access to record a specific piece of data then a paper-based or application-based method of recording is likely to be preferable to a system that is online available via a computer or the Internet, as these may mean data recording is missed when these are not available.

Systems that allow users to collect information without providing any means through which they can reflect on data, or gain self-knowledge are not included in this class of system. For example, Foursquare<sup>3</sup> allows users to share their location but does not provide a context within which this data can be used for self-improvement or -

---

<sup>1</sup><https://gyroscop.pe/>

<sup>2</sup><https://exist.io/>

<sup>3</sup><https://foursquare.com/>

understanding. Where Do You Go<sup>4</sup>, however, allows users to visualise their Foursquare check-ins to better understand their movements. Thus, Where Do You Go would be considered a PI system, whereas Foursquare by itself would not be. The definition of PI tools/systems put forth by I. Li et al. (2010) (see Chapter 1, 1.1) is used throughout this research.

## 2.3 Novel Sensor Technologies

New sensing technologies are becoming available at a rapid rate, providing methods for advanced sensing and analysis in a variety of domains. These sensing technologies fulfil essential roles in the the domains of medicine, earth sciences, natural sciences, sports, critical control systems, media and entertainment, to name a few. Sensing technologies are used to gain a better understanding of the natural world (Ruiz-Garcia, Lunadei, Barreiro, & Robla, 2009), detect medical issues for patients (Falck, Baldus, Espina, & Klabunde, 2007), provide advanced warning systems in the case of natural disasters (Mousa, Zhang, & Claudel, 2016), and enhance entertainment; from detection systems for sports (Chi, 2008) to increasing a sense of realism in animated films and TV by capturing actor data more accurately.

Some of these new sensing technologies allow the user to track things that previously may have been difficult to do manually, or enable the detection of measurements that cannot easily be recorded by a person by themselves. These technologies offer a new potential for providing insights about a person, which may have previously been deemed accessible only to trained science or medical professionals, or affordable by large commercial enterprises. As novel sensing technologies become available to consumers, they offer the ability for the ‘average’ technology user to take ownership of collecting and analysing data themselves for their own purposes.

Often times personal sensing technologies are developed with the aim of launching a commercially successful business. However, the veracity of the data that is captured - the degree to which it is accurate, precise and trusted - is not always backed up by scientific research. In some cases technologies may have been subjected to in-house testing to ensure that the technology works in a sufficiently acceptable manner to be accepted by consumers, but largely the technologies have varying levels of reputable, repeatable and certifiable research studies to back them up. Gillinov et al. (2017) showed that wearable devices used to detect heart rate vary between wearable devices and even between activities. The authors suggest that manual palpation or the use of an ECG as

---

<sup>4</sup><http://www.wheredoyougo.net/>

preferred ways of evaluating heart rate. Further studies have presented similar results that suggest that accuracy of wrist-worn devices being less than that of chest-worn sensors (Etiwy et al., 2019). Even technologies which have been around for several years, have garnered mainstream adoption and might be considered commercially successful, such as Fitbit, have had their accuracy and reliability questioned (Sawyer, 2019).

The following sections highlight several characterisations of technologies that that we consider as typifying a ‘novel sensor technology’ within the context of this research.

### **2.3.1 Data & Sensor Availability Increases, Cost Decreases**

Sensing devices are now pervasive. Generally speaking, there has been a rapid increase in sensor production during recent decades, but the current wave of ‘Internet of Things’ (IoT) adoption is in large part responsible for increasing the ways in which we can quantify how we interact with the world (Swan, 2012). Decreasing sensor size allows more sensors to be packed into a variety of devices. Typically one can find a modern smartphone to contain an accelerometer which can track a phone’s movement, a gyroscope which can measure a phone’s orientation, magnetometers which can measure magnetic fields and thus can be used as compasses, ambient light sensors to adjust screen brightness, GPS which can be used to determine location, proximity sensors which can sense how close a phone is to one’s head. They may also contain sensing technologies in a broader sense, including visual sensors such as cameras, sound sensors such as microphones, fingerprint sensors, touch sensors. Additionally, sensor data may be combined through a process of ‘data fusion’ (Hall & Llinas, 1997), or appropriated for new sensing applications. For example, a smartphone may make use of an accelerometer to provide an indication of physical activity based on phone movement, which can be further enhanced with an ability to detect heart rate using a combination a phone’s light and camera (Bolkhovskiy, Scully, & Chon, 2012). The number of sensors produced annually is increasing and it is further predicted that we will pass 1 trillion sensors in the near future (Bogue, 2014).

The lowering cost of available sensing technologies is likely to make sensing technologies more accessible to users, in much the same way that decreasing costs in computing and mobile phones increased their availability and uptake. The miniaturisation and decreasing cost of producing and manufacturing sensing technologies will make them more adoptable. Emerging technologies such as smartwatches are becoming increasingly accessible and familiar to people and the majority provide some sensing capabilities that could be used to support data collection for self-tracking and self-reflection.

This miniaturisation, lowering cost and increasing availability of sensors has a particular impact within the domain of personal electronics, in particular in relation to the growth in consumer wearable devices, as discussed in the following section.

### 2.3.2 Sensors are Wearable

The current ubiquity of wearable technologies is rooted in the progression of research in the areas of wearable computing. Although the inclusion of technology in things that can be worn begins much earlier, such as wrist-watch calculators developed in the 1970s, wearable computing research has roots in the 1980s and 1990s research that includes developments by Steve Mann (Mann, 2018). Mann’s initial experiments with wearable computing led to the design of the first wearable general-purpose computing system was developed with applications in augmented and/or virtual reality, including the development of head-mounted displays (Mann, 1997). Other developments by Mann includes the EyeTap system which was capable of recording data from the wearer’s eye’s perspective (Mann, Fung, & Moncrieff, 1999). Further research by Mann in this field led to the development of a wearable computing system that was capable of recording video of the wearer’s surrounding, Mann coined this area as ‘sousveillance’. Mann also describes his work in relation to other works such as the development of Cyborglogs, whereby a cyborg may “generate a lifelong logfile for personal experience capture” (Mann, 2004b) with EyeTap research also focusing in areas including continuous logging of personal experiences (Mann, 2004a).

During the same period as Mann’s development and progression of research from wearable computing to continuous logging of personal experiences, advancements were also seen in other domains which led to the development of modern ‘activity trackers’. The first commercially available wearable/personal heart rate monitor was released by Polar in 1982, the PE2000 (Laukkanen & Virtanen, 1998). In 1984 they released the Polar Sport Tester PE3000, the first wearable heart rate monitor that provided a screen on which the wearer could view their heart rate (Laukkanen & Virtanen, 1998). This device was aimed at providing athletes with insights into their training. Further developments followed, with technologies capable of recording physiological signals used in the domains of medical wearables and affective wearables. Picard and Healey (1997) describes devices which allow for the continuous monitoring of physiological signals using sensors embedded in shoes, earrings, or worn on the hand in the case of affective wearables (those which attempt to monitor a person’s affective state e.g. stress, fear, relaxation), whilst medical wearables are more concerned with providing medical devices that can be used in ambulatory situations (i.e. situations where the wearer may



be moving such as walking or partaking in a physical activity). The authors describe this advancement in wearable computing, and its overlap with ‘medical wearables’ as providing “an unprecedented opportunity to ‘get to know’ a person” (Picard & Healey, 1997).

It is the combination of this research that leads to the advent of what we today would consider ‘activity trackers’, such as Fitbit that allow the wearer to record a combination of physiological signals, and smartwatches that have physical activity tracking functionality built in.

The range of commercially available sensing technologies is vast and growing rapidly, these range from clip-on and strap-on sensors such as those manufactured by Polar<sup>5</sup>, to watches<sup>6</sup> and wristbands<sup>7</sup>, and rings<sup>8</sup>. Seneviratne et al. (2017) provides a classification of wearable sensing technologies that are currently available and their intended positioning on the body, shown in Figure 2-3. The devices mentioned here present a snapshot of the types of technologies available at the time their paper was published (2017), which typify ‘wearable technologies’ and provides a basis for identifying technologies that may become increasingly pervasive in the near future.

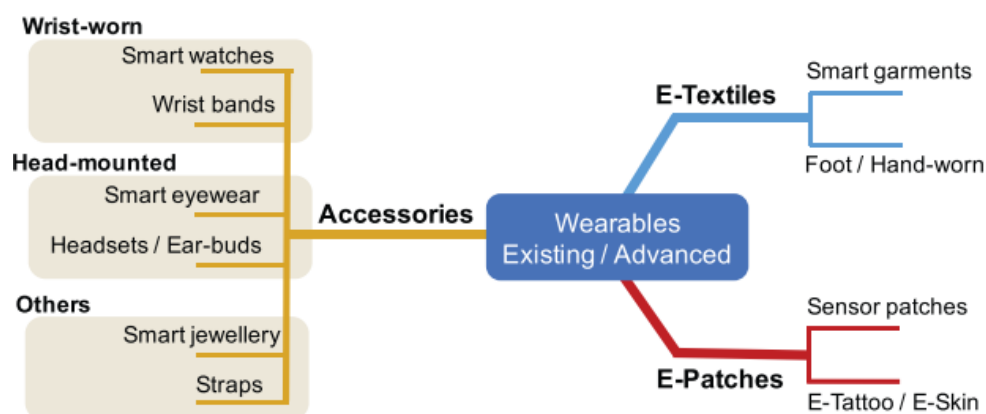


Figure 2-3: Classification of wearable devices (Seneviratne et al., 2017)

A systematic review of 463 published articles related to wearable activity trackers highlights the increased focus on wearables in research. The review found that the number of published studies involving wearable activity trackers had increased almost 25-fold from 2013 to 2017, from 8 studies conducted in 2013 to 199 in 2017 (Shin,

<sup>5</sup><https://www.polar.com/uk-en/products/heart-rate-sensors>

<sup>6</sup><https://www.fitbit.com/us/products/smartwatches>

<sup>7</sup><https://www.mi.com/uk/mi-smart-band-4/>

<sup>8</sup><https://www.ouraring.com/>

Jarrahi, et al., 2019).

Currently the consumer wearable sector is dominated by wearable technologies that relate to supporting active lifestyles and as such are focused on a wide variety of sports and/or activity tracking. A 2015 review of wearable sensors for human activity monitoring found that the main consideration that was holding back development was cost, but that there is an expanding market, and this is expected to grow in the future (Mukhopadhyay, 2015). More recently Aroganam, Manivannan, and Harrison (2019) performed a review of wearable sensors used within sports, excluding professional sports, with the authors highlighting the considerable consumer market share of wearable devices, as can be seen in Figure 2-4.

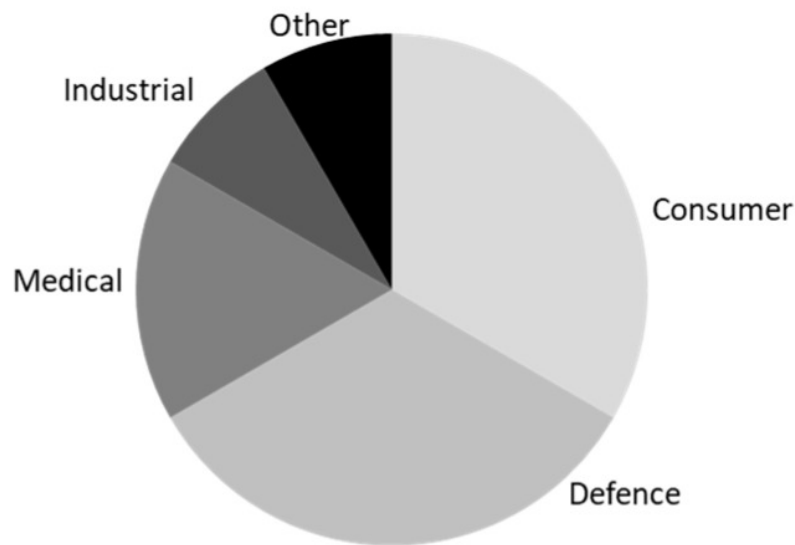


Figure 2-4: Wearables market share (Aroganam et al., 2019)

Wearable technologies also have a growing market in the areas of meditation and mental health, with Muse<sup>9</sup> and Emotiv<sup>10</sup>, two EEG headset manufacturers, both providing consumer-aimed products. Google is also weaving smart-tech into clothing with Jacquard<sup>11</sup>.

It is through this progression and combination of research areas that we reach our current research focus on the use of wearable technologies in Personal Informatics and the Quantified Self movement. The type of technologies we consider in the scope of this research is based on this, as well as the growing availability and increasing wearability

---

<sup>9</sup>[www.choosemuse.com](http://www.choosemuse.com)

<sup>10</sup>[www.emotiv.com](http://www.emotiv.com)

<sup>11</sup><https://atap.google.com/jacquard/>

of consumer sensing technologies.

### 2.3.3 Data Collection is Automated

Manual self-tracking is the process of a person manually recording data. This may be in the form of pen and paper style journals or note taking, but could also be done digitally using spreadsheets, word processors or an app specifically designed for the purpose of making manual entries. Examples of the types of data that are often tracked manually include food diaries (Burke, Wang, & Sevick, 2011), journals (Pomfret & Medford, 2007), and mood logs (Matthews, Doherty, Sharry, & Fitzpatrick, 2008).

The complexity and depth of data to be tracked varies according to context and the needs of the individual. As I write this thesis I am manually tracking my word count in a spreadsheet on a daily basis so that I can track my progress and determine my levels of productivity. Although this is a form of *digital* tracking, an important distinction is that data collection is not necessarily automatic by virtue of being digital.

In contrast to manual tracking, automated tracking allows for people to record data with minimal interaction with an analogue or digital system. Examples include step counts which can be inferred by mobile phones sensors or wearable activity trackers, heart rates measured by wrist- or chest-worn sensors, or sleep tracking inferred via phone usage, bed sensors, or wearable technologies. Emerging sensors, such as the previously mentioned, Feel<sup>12</sup> and Upmood<sup>13</sup> wearables, claim to automate the collection of mood data for the wearer.

Although self-tracking often incorporates manual recording of data, this places a significant reporting burden on the user (I. Li, Dey, & Forlizzi, 2012). Hence, there is an increasing shift towards methods of automation, both for data collection, as well as for data analysis, i.e. making people more aware about the data and the insights that it can provide (S. L. Jones & Kelly, 2018). Automated data analysis, in this context, refers to the automation of providing people with an understanding of the data, which they may not be able to uncover by themselves manually. However, automated data collection and automated data analysis are not mutually inclusive. The provision of a technology capable of automatically tracking data does not always come provided with a means of automatically analysing its data.

---

<sup>12</sup><https://www.myfeel.co/>

<sup>13</sup><https://upmood.com/>

### 2.3.4 Data Collection is Unobtrusive

The miniaturisation of sensing technologies leads to a decrease in the obtrusiveness associated with technologies used for collecting data. Zheng et al. (2014) highlights the proliferation and possibilities of unobtrusive sensing within the context of health informatics. For example, their research highlights that in a medical/health informatics context “miniaturization and unobtrusiveness can enhance the comfortness [*sic*] of using wearable devices, and thus increasing the compliance for long-term and continuous monitoring” (Zheng et al., 2014). Furthermore, this enables “personalizing the design of the devices”, in terms of “personalizing the sensor calibration, disease detection, medicine, and treatment” (Zheng et al., 2014). Although health informatics is concerned with the collection and analysis of continuous medical or biomedical data, there is considerable overlap in the types of technologies used in PI settings.

Lupton (2014) describe the blurring boundary between the body and the digital as being “most overtly the case of such technologies as heart pacemakers, insulin pumps and cochlear implants, all of which are inserted into the body in unobtrusive ways” (Lupton, 2014). The blurring and unobtrusive nature of medical devices as part of the human suggests a possible next step for PI devices too.

Lawson et al. (2013) describes their novel sleep sensing application as an unobtrusive means of collecting sleep data. Their work builds on previous research by Choe, Con-solvo, Watson, and Kientz (2011) that highlights people’s requirements for unobtrusive technologies for sleep monitoring. Research also presents opportunities to measure values unobtrusively, such as mental states using mobile phone trace analysis (Canzian & Musolesi, 2015). The design objective of producing an ‘unobtrusive’ technology may also be considered in terms of mechanisms for feedback to PI system users, such as described in work by Yu and Ingalls (2011), who sought to provide people with a way to track and better understand their web browsing habits via a post-hoc data visualisation tool rather than using alarm-like notifications as a means of drawing attention when attempting to increase productivity and making people aware of their habits.

### 2.3.5 Data Collection & Analysis is Personal

Given the previous characteristic of wearability, the personal nature of the data and its analysis increases, as sensing technologies become more connected and integral to the person. In a PI setting, data is potentially highly meaningful to the person who collected it. Although there may be a third-party company involved (e.g. processing data, or providing storage and analysis capabilities), the ways in which it is analysed,

collected and interpreted is of concern to the person; a defining characteristic of the technologies considered as part of this research. This feature is implicit in the nature of PI and the QS movement. The personal nature of the data collection and reflection process empowers peoples' self-determination and responsibility (Gimpel, Nißen, & Görlitz, 2013).

Sensor technologies, including wearables, are spread across a variety of domains including medicine, sports and athletics. In these domains, the intention may be that a domain-expert reviews and analyses the data on the wearer's behalf. Given the nature of how this data is collected and reviewed, we would not consider these within the scope of this research. Some technologies may be considered in both domains as their data could still be accessed and analysed by the wearer of the device.

### **2.3.6 Data is Multifaceted**

With the proliferation of sensing technologies data can be combined with other sensors' data streams to provide new insights that would not have been discovered using un-faceted data analysis. For example, data streams for sleep and physical activity could be combined in a correlational analysis (S. Jones & Kelly, 2016) to uncover what impact, if any, someone's physical activity has on their sleep patterns, or vice versa.

The multifaceted nature of PI systems is described by I. Li et al. (2010). An exploration of tools for combining multiple data sources via visual 'cuts' was explored by Epstein et al. (2014), and an understanding of multifaceted data analysis systems is explored in S. Jones and Kelly (2016). Research by I. Li et al. (2012) highlights the inclusion of contextual information (which gives additional meaning to data), as a facet, as beneficial for self-reflection in the long-term.

Combining multiple data sources allows a richer exploration of data, however, the standards for recording different data may vary depending on the technology and its design. Therefore collected data may be structured, semi-structured or unstructured. This poses challenges for designing effective interfaces that allow insights to be extracted from multifaceted datasets, the likes of which may be common in a PI context.

### **2.3.7 Transparency of Data Collection is Unclear**

Transparency, as defined here, relates to the ability to see through or into the inner-workings of something. In this thesis, transparency is discussed in relation to the inner-workings of a sensing technology or digital device. Transparency in this sense has some overlap with research in other computational science related areas such as

artificial intelligence, machine learning, big data and digital/data privacy. Although an open and clear explanation of how a company may intend to use a person's data is important, the typification of a novel sensing technology's transparency in this context is referring the ability of user to be able to determine how a sensing technology has arrived at a presented piece, or pieces, of data rather than how companies may use a person's data. With modern sensing technologies the data presented to the users may be abstracted away from the raw values being recorded by the sensors, and presented in a more 'user-friendly' or easily absorbed piece of information. The provided values may be the product of some internally calibrated algorithm or some proprietary algorithm leaving users with little understanding of how the the values are actually calculated. For example, a sensor in a pedometer may use a combination of movement sensing and machine learning to detect and track the activity taking place (e.g. walking, cycling, running) but the how and why of how the system has determine this is not always obvious to people, or even available to be inspected should they wish. The opposite of transparency is sometimes termed as 'black box' whereby one cannot be certain of the inner-workings of the thing contained within 'the box'. As with most modern electronic systems there is usually some kind of inner algorithm, whether it be sophisticated or simple, to which most users do not have access to or the ability to see.

It might be assumed that people have some understanding of what values are represented by "heart rate" or "breathing rate" but how the sensing technologies come up with these values for someone is not always transparent. Yang, Shin, Newman, and Ackerman (2015) conducted an analysis of 600 product reviews and found that people's perceptions of the product and the ability determine if a device is accurate may be affected by the transparency of a device (or its manufacturer) in providing sufficient information about how a device has determined the data that is being presented to users. Such a misunderstanding might lead a user to believe that a system uses the time of day to determine sleep patterns rather than a lack of device motion, or that the products ability to correctly calculate how many calorie have been burnt is incorrect, for example when a users is aware that they have not moved around much on a particular day but the device shows a higher than expected value.

In a research study which "sought to explore users' direct reactions to a transparent PI system that interprets emotions that users express in written text" (Springer & Whittaker, 2019), the authors increased the transparency of the system by showing how different words' negative or positive weightings resulted in a specific emotion being detected. A graph of the overall detected emotion was presented to participants and in the transparent condition participants were able to see colour-coded highlighting

that showed whether a word had been categorised as positive or negative. In the non-transparent condition participants saw no highlighting and were only able to see the graph. An important insight from this paper is that users may overestimate a systems capabilities if they are less transparent, as their qualitative findings suggest that “users who know less about the working of the system, seem to ascribe more advanced abilities to it” (Springer & Whittaker, 2019). The authors also found that before using the system personally participants predicted that the transparent version of the application was more accurate. In the second part of the study, the authors found that “users initially form simple working hypotheses about system operation. Users seem to engage with transparency first by operating with these simple hypotheses and only scrutinizing these when their expectations aren’t met” (Springer & Whittaker, 2019).

Respondents in Liang, Nagata, Martell, and Nishimura’s (2016) online survey study were asked about their mindset, toolset and skillset as related to mHealth and self-care highlights technological transparency as an issue. The authors suggest that improving technology literacy is one challenge that should be addressed in the future as ““not knowing [*sic*] the mechanism behind the measurement” made it hard for users to judge the performance and accuracy of the technologies, and thus affected users’ trust and adoption” (Liang et al., 2016).

### **2.3.8 Data May Have Varying Levels Trustworthiness**

As mentioned, transparency of how values are measured impacts people’s trust in technologies and people may not be able to accurately determine the level to which they should trust the values produced by these technologies. The majority of currently available smartwatches capable of measuring heart rate have no medical certification for their reliability of the readings they produce. However, this does not necessarily mean that they do not produce values that can provide useful for actionable insights.

Research conducted by Jaimes, Murray, and Raij (2013) looks at the ways in which we might enable increased trust in PI systems in relation to the uncertainty, noise and measurement error in the information that is displayed in PI tools. One factor that the authors suggest might result in a reduction of trust is when a person’s memory of an event differs from what is presented by the system. Jaimes et al. (2013) also draws attention to work by Lim (2010) which suggests that a person’s inability to trust a device may lead to a loss of interest, and work by Dzindolet, Peterson, Pomranky, Pierce, and Beck (2003) and Kulesza (2012) which suggests data errors can also lead to a reduction in trust which may eventually lead to abandonment of the technology.

Technological trust plays an important role in ensuring that people are able to achieve their goals, for example increased physical activity, as peoples' perceived trust in a device directly impacts whether they will believe the data presented by tracking devices (Michaelis et al., 2016). However, as also highlighted by Michaelis et al. (2016), it has been suggested that a device's reliability rather than it's validity is an important factor for use (J. Meyer, Fortmann, Wasmann, & Heuten, 2015).

An important thing to note, as highlighted by Hollis, Pekurovsky, Wu, and Whittaker (2018), in their study on deference to algorithmic emotion detection, is that increased algorithm transparency may not actually increase user understanding and there may be an unjustified deference to believe what an algorithm has presented as being the truth.

It is clear from the above that both technological trust and transparency are two related concepts and that having more of one may increase the other, although there is research that also argues that it might not. Therefore there is potentially a double-edged sword, whereby as a technology boasts an increasing accuracy, this could itself lead to a situation people rely on the accuracy attested, which could in itself lead to an over-reliance on technology and potentially detrimental deference to that technology, which might be prevented if the algorithm can be scrutinised via transparency - at least by those who want to, or can do.

### **2.3.9 Data May Lack Familiarity**

New data sources present people with obstacles in first becoming familiar with the data and what specific values may mean. The person may not know what values are acceptable, i.e. is the sensor actually detecting and presenting values that are within the range that one would consider 'acceptable'; people may not be aware of what values they should consider 'good', 'bad', 'neutral'. Is the data source actually providing values that actually make sense? Hypothetically speaking if someone had no idea about the heart, heart rate or pulses, and put on a sensor that was meant to detect their heart rate but the sensor always presented a value of 10, how would the person know that the value is incorrect or that they may be in severe distress? Leaving aside the unlikely scenario that someone with a heart rate of 10 would not actually be able to read the values, without some sort of domain knowledge or prior experience it is very difficult to know for certain what any presented value means, or that perhaps the sensor is not working correctly and the defect lays in the sensor and not the signal being measured.

People may lack prior experience of, or familiarity with, the types of data that are



collected using the sensing technology. Some unfamiliarity with PI systems is in part due to not having previous experience with self-tracking and PI (c.f. (Rapp & Cena, 2014; I. Li et al., 2010)). The familiarity focused on in this characteristic is due to a person’s lack of familiarity with the data source rather than the practice of self-tracking. Research proposes that new PI users may not be inclined to spend as much time becoming familiar with how data is tracked, the visualisation of quantified data that PI systems provide, as well as being unsure about the benefit a particular data source may provide (Rapp & Cena, 2016).

### 2.3.10 Data has Volume, Variety, Velocity & Veracity

Some of the previously described characteristics share commonalities with the 4 Vs of ‘Big Data’: volume, veracity, velocity and veracity described as (Miele & Shockley, 2013):

- **Volume:** The amount of data. . . . volume refers to the mass quantities of data that organizations are trying to harness to improve decision-making across the enterprise.
- **Variety:** Different types of data and data sources. Variety is about managing the complexity of multiple data types, including structured, semi-structured and unstructured data.
- **Velocity:** Data in motion. The speed at which data is created, processed and analyzed continues to accelerate.
- **Veracity:** Data uncertainty. Veracity refers to the level of reliability associated with certain types of data.

Although volume as defined above is being defined in relation to business this could be reconsidered in the context of a person, given that previous characteristics of increasing availability and access to personal data, as: the quantities of data that *people* are trying to harness to improve decision-making across *their lives*. In their research into an ‘Augmented Smart Coach Based on Quantified Holistic Self’, Yoon, Doh, Yi, and Woo (2014) describe these in relation to the quantified self as *personal big data*.

The volume, variety and veracity of data that people collect are likely to increase as has previously discussed in the previous subsections on increasing availability of sensing technologies and the transparency and trustworthiness of the collected data. The velocity of data considers the speed at which the data is recorded and becomes available. Sensors provide the ability to increase the rate and speed at which data

is collected, which in turn increases that amount of available data also. Katz, Price, Holland, and Dalton (2018) conducted a study of people living with diabetes who self-tracked, such people may have to take blood samples and record blood glucose levels at intervals throughout the day. The authors highlight that people have issues interpreting this data and that “there is a need to re-consider how to help users draw value from real and often noisy diabetes data”. As an example situation, these issues may be compounded even further with advances in technologies such as continuous blood glucose monitoring technologies in the future which provide larger data volumes at much greater velocity by providing a continuous reading of blood glucose levels. These are currently available as medical devices but Dexcom is developing PI-style tracking devices<sup>14</sup>, and news reports suggest that Apple are interested in trying to build this type of sensing technology into their watches, as suggested by a patent submission reported by Healio Endocrinetoday (2020) and Wearable Technologies (2019).

## 2.4 Sensors Used in This Research

Sensing technologies allow for the automated collection of data, reducing the burden of manual data logging for users. While it is possible for someone to spend time manually logging every step they take or continuously measuring their heart rate, it is highly impractical if they wish to do anything other than continuously measure. Another practical use for sensors is to allow the measurement of things that are not as easy to (self-)measure manually, for example, posture detection sensors<sup>15</sup>.

A commonly used sensing technology are step counters, or pedometers, which allow the automated counting of the number of steps the wearer takes. This is done using a sensor that is capable of detecting movement of the device itself and software that is calibrated to convert this sensed movement into an equivalent number of steps. Modern devices capable of tracking steps counts make use of an accelerometer. An accelerometer is a sensor capable of measuring acceleration within a three-dimensional plane, by measuring these movements it is possible to determine an estimated step count (Mladenov & Mock, 2009). With advances in machine learning techniques these measurements can be made more reliable and people have been able to further expand the more ‘simplistic’ capabilities to also include detection of specific activities, e.g. walking, running, cycling (Ravi, Dandekar, Mysore, & Littman, 2005). Such sensors have varying levels of accuracy (Huang, Xu, Yu, & Shull, 2016; Thorup et al., 2017) when measuring true walking. One further problem they can face is that due to the

---

<sup>14</sup><https://www.dexcom.com/en-GB>

<sup>15</sup><https://www.uprightpose.com/en-gb/>

inability of the device to determine if the sensor has sensed a ‘real’ step versus incidental movement or shaking of the device.

In the following subsections we describe the three types of sensing technologies that we use in our studies. The described sensors are built into the devices which participants used during our studies.

#### **2.4.1 Heart Rate Sensor - Fitbit**

Heart rate detection is becoming increasingly ubiquitous. A variety of smartwatches (e.g. Fitbit, Apple Watch) include the ability to measure the wearer’s heart rate. Some phones (e.g. Samsung Note, Samsung Galaxy) also include the ability to record the user’s heart rate on demand. This is done using a technique known as photoplethysmography. Photoplethysmography is a process by which light is shone on, for example, the surface of the skin and measurements of the changes in this light can be used to indicate heart rate (Allen, 2007). Aside from smartphones and smartwatches, more traditional medical devices also include this ability, such as at-home blood pressure monitors without the use of photoplethysmography. For example, Omron<sup>16</sup> produces blood pressure monitors that also detect heart rate. Omron also provide information regarding whether their monitor has been clinically validated, backed by peer-reviewed research (see Asmar, Khabouth, Topouchian, El Feghali, & Mattar, 2010). This makes Omron the exception rather than the rule, compared to the Fitbit devices used in our research. As mentioned previously, the accuracy of Fitbit has been questioned in the past both in terms of step count accuracy (Huang et al., 2016; Thorup et al., 2017), and heart rate detection (Sawyer, 2019). The Fitbit device used by participants in our studies can be seen in Figure 2-5.

#### **2.4.2 Breathing Rate Sensor - Spire**

Many wearable sensors capable of breathing rate detection come in the form of a band that the wearer straps around their chest (Nikolic-Popovic & Goubran, 2011), however, this form factor is becoming smaller. Spire<sup>17</sup> was able to fit the required sensing technology into a pebble-shaped device that the wearer could attach to their belt, or attach to clothing (see Figure 2-6). When attached the sensor in the pebble is able to detect changes in the the devices movements which are translated into a number of breaths per minute.

---

<sup>16</sup><https://www.omron-healthcare.co.uk/>

<sup>17</sup><https://www.spirehealth.com/>



Figure 2-5: Fitbit Versa



Figure 2-6: Spire Pebble

### 2.4.3 Electroencephalography - NeuroSky MindWave

Electroencephalography (EEG) is an example of a near-future technology that is becoming more portable, wearable and is now seeing input from big technology companies eager to cement themselves as the founders of consumer brain-computer interfaces (BCIs). EEG sensors allow for the recording of the electrical signal of the wearer's brain. Our reason for choosing this sensor as a near-future technology probe for the research conducted as part of this thesis is described in Section 2.5. Further background information regarding BCIs and EEG is presented in section 2.6. The NeuroSky MindWave used by participants in our studies can be seen in Figure 2-7.



Figure 2-7: NeuroSky MindWave headset

## 2.5 Technology Probe Sensor Choice

In this thesis we focus on investigating the design and use of sensing technologies and PI systems that share the characteristics described in the Section 2.3. The data collection is *Automated*; the collected data is *Personal*; the sensor technology is readily *Available* and can be obtained at low *Cost* through the use of off-the-shelf *Unobtrusive* and *Wearable*, technology. The sensing technologies produce a significant *Volume & Variety* of data that can be collected. The sensing technologies provide a variety of data (i.e. data is *Multifaceted*); and people may lack *Familiarity* with the data being collected; the *Trustworthiness* of data may vary or be difficult to determine, and the *Transparency* of how data is produced and processed may vary.

As a way of exploring this, this research uses a consumer-grade brain-computer interface headset as an example of a near-future sensing technology. This is a device that people might consider as a technology that may be used in the future but that is still in its infancy such that people would likely have little-to-no previous experience with the technology. A commercial BCI typifies an emerging sensing technology as previously described due to the following characteristics: it is personal as it produces data about a person's brain, it is wearable as it is worn on the person's head, the type of data it produces is likely to be unfamiliar to a majority of participants. The technology is currently commercially available but none currently have medical certification and the transparency of how EEG values are calculated are likely unclear to people. EEG sensors tend to have a lot of noise associated with the signal acquisition due to placement of the device. The values produced can be integrated into data-rich and

multifaceted PI systems.

A Fitbit smartwatch and a Spire breathing rate monitor, are also used in different parts of the research described in this thesis. These technologies provide both a way to enable multifaceted tracking by providing participants in studies with a secondary source of data, and less abstract sources of data. Only the heart rate data is recorded from the Fitbit watch (although nothing prevented participants from viewing the other data if they so chose) and only breathing rate data is recorded by the Spire stone. Therefore also limiting the variety of data these sensors are used for, making them more ‘simple’ sensors.

## 2.6 Electroencephalography and Brain-Computer Interfaces

Brain-computer interfaces, the generic term used for electroencephalography technologies when used in combination with computing technology, allows for the acquisition and recording of the minute electrical voltages produced as a by-product of brain activity, often referred to as ‘brainwaves’. The recording of brain wave activity has been around for about 100 years and has applications in medicine and currently there are several commercially available examples of BCIs: Muse, NeuroSky<sup>18</sup> and Emotiv are three of the most well-known manufacturers currently but several other companies are also involved in the development and sale of BCIs as devices intended for use by the general public. These devices provide the wearer with access to data produced by their brain that they can potentially interpret for several purposes. However, the purposes for which that data can be used or should be used is still being researched, in particular for use by ‘consumers’ rather in a clinical setting.

As the raw signal recorded by these devices is generally expressed as a voltage, the raw values need to be interpreted in some manner either directly by the manufacturer of the device or by interpretation of the raw values by a third-party. Thus both the raw signals acquired via the device and the abstracted values are likely to be unfamiliar to the user.

The following sections discuss the background of electroencephalograph, brain-computer interfaces (BCIs), consumer BCIs and the use of BCIs within the context of PI.

---

<sup>18</sup>[www.neurosky.com](http://www.neurosky.com)

### 2.6.1 Electroencephalography

Electroencephalography (EEG) is the method by which brainwaves are recorded via electrodes placed on a subject's scalp. Hans Berger invented the electroencephalogram in the 1920s, he also recorded the first human brainwaves during the same period (Haas, 2003). The brain signal of a subject can be recorded due to the electrical action potentials generated by brain cells. However, as the signals picked up by the electrodes are very weak an amplifier is required to boost the signal so that it can be interpreted.

EEG differs from other means of recording brain activity such as Electrocorticography (ECoG), also known as intracranial electroencephalography (iEEG), as it does not require direct contact with the brain (Figure 2-8).

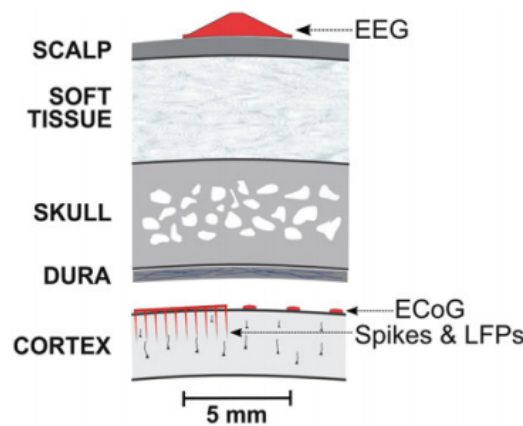


Figure 2-8: Recording sites for electrophysical signals (Wolpaw & Wolpaw, 2012)

The recording and evaluation of EEG data are routinely used in clinical practice (Figure 2-9) for detecting brain anomalies (Ocak, 2009) and there is growing research into the use of EEG for controlling assistive technologies, e.g. prostheses (Müller-Putz & Pfurtscheller, 2008).

### 2.6.2 Brain-Computer Interfaces

Brain-computer interfaces (BCI) build upon the electroencephalogram technology and enable the processing of EEG data via a computer to allow some form of interaction to be achieved. The basic design of a BCI system (Figure 2-10) consists of acquiring the EEG signal, performing feature extraction on the signal and then translating those features in a way that allows an action to be performed within the system based on these features.



Figure 2-9: Clinical EEG recording (NHS Choices, n.d.)

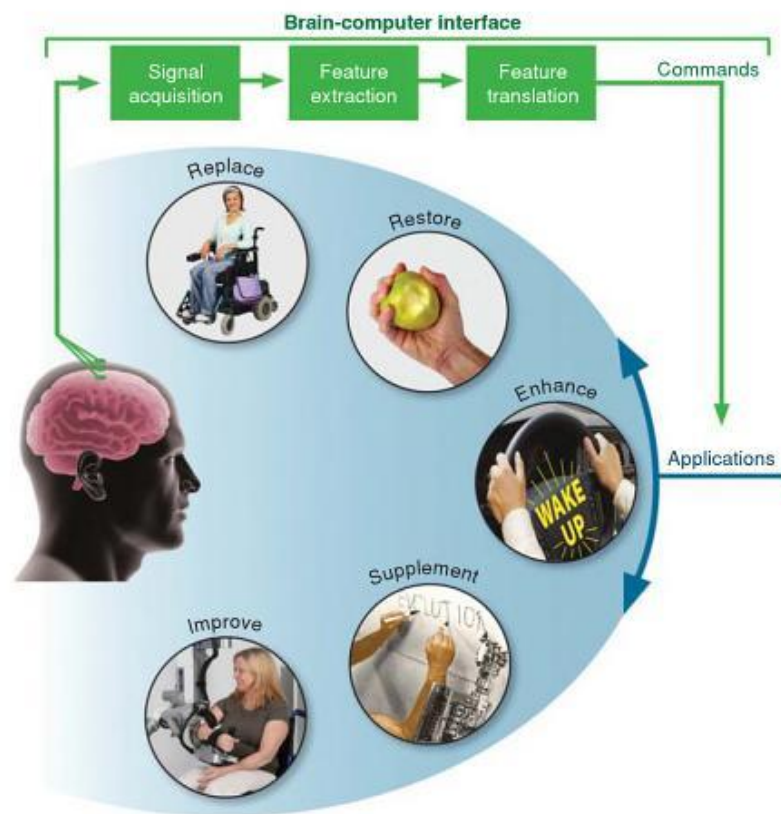


Figure 2-10: The basic control and design of a BCI system (Wolpaw & Wolpaw, 2012).

The first research into the uses of brain-computer interfaces was in 1973 by Vidal (1973). A large focus of BCI research has been on using BCI devices to help those



with disabilities. Applications include prosthesis control (Müller-Putz & Pfurtscheller, 2008) for amputees, BCI spellers (Yin et al., 2013) to help, for example, people with varying types of paralysis. However, these studies and applications often make use of medical-grade EEG recording equipment. Medical-grade EEG devices can provide higher dimensions of brainwave data but as a trade-off they tend to be large, cumbersome and require quite a bit of set up (see Figure 2-11), this is likely to be a negative factor for consumer uptake.

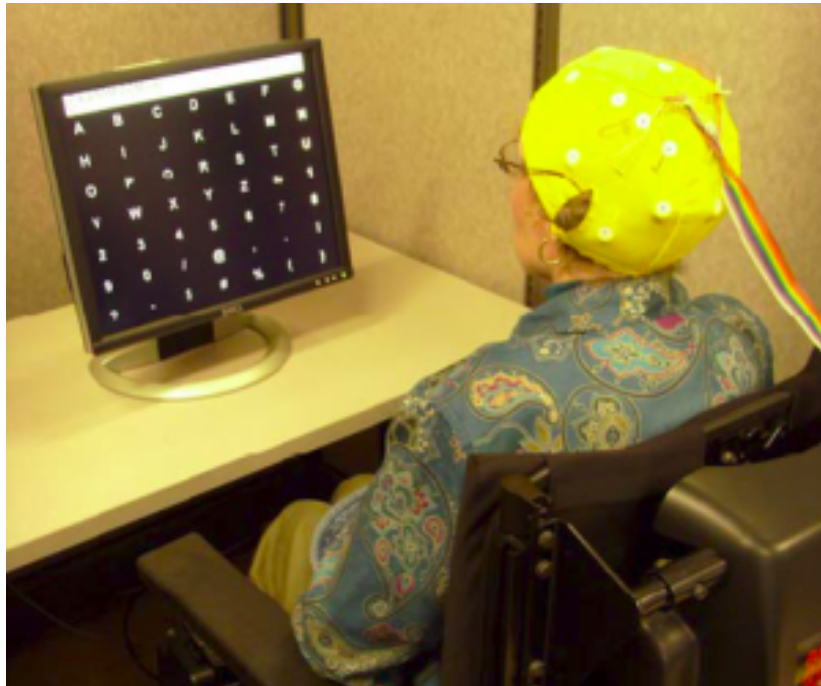


Figure 2-11: Medical-grade BCI Speller system (Y. Li et al., 2014).

### 2.6.3 Consumer-Grade BCI

The emergence of low-cost, consumer-grade EEG/BCI headsets from companies such as NeuroSky and Emotiv, enables EEG recording devices to be obtained at a reasonable cost for personal use. The NeuroSky MindWave headset (Figure 2-12a) provides a single dry electrode positioned just above the user's eyebrow. The headset's data can be accessed via a software development kit (SDK). Raw EEG data can be sampled at 512Hz. Values for each of the five commonly used EEG wave bands (Alpha, Beta, Gamma, Delta, and Theta) can be accessed at a rate of 1Hz. The SDK also provides access to two in-built algorithms for detecting attention and meditation values; called eSense values, as well as blink detection. In comparison, the Emotiv EPOC (Figure 2-12b) provides 14 sensors, which need to be saturated in saline solution before use. The

EPOC provides raw EEG at 256 samples per second, per electrode. The EPOC is also capable of providing gyroscope data, as well as blink detection. Both the MindWave Mobile and EPOC are wireless, with the ability to communicate data via Bluetooth.



(a) NeuroSky MindWave headset (Neurosky, 2017)



(b) Emotiv EPOC (EMOTIV, 2017).

Figure 2-12: Consumer-grade BCI headsets

The MindWave headset has been used in a variety of studies to ascertain its efficacy (e.g., Maskeliunas, Damasevicius, Martisius, & Vasiljevas, 2016; Rebolledo-Mendez et al., 2009). Crowley, Sliney, Pitt, and Murphy (2010) used the MindWave to detect stress by monitoring the meditation and attention eSense values of participants as they performed the Stroop Colour-Word Interference Test and a “stress-inducing” puzzle game, The Towers of Hanoi. They found that the data from the headset could indicate whether a participant underwent a change in their emotional state. However, it was not possible to accurately determine an exact instance at which the change occurred, rather the data was able to provide an overall trend of the participant’s emotion. A similar study by Chee-Keong Alfred and Chong Chia (2015), used the Stroop Test to induce a stress response in participants, but their research used the raw EEG data, rather than NeuroSky’s eSense values or power band values, and additionally looked at potential machine learning classifiers that may improve their ability to detect stress. The results demonstrated that the use of machine learning classifiers enabled them to predict the self-perceived stress state (stressed or non-stressed) of participants from their EEG data, with an accuracy of 72% (Chee-Keong Alfred & Chong Chia, 2015).

## 2.7 Brain-Computer Interfaces for Personal Informatics

Most studies of consumer-grade BCI devices focus primarily on their use as input/control devices, e.g. in gaming (Van De Laar, Gürkök, Plass-Oude Bos, Poel,

& Nijholt, 2013), or as tools for improving focus and engaging with mindfulness exercises (Bhayee et al., 2016), despite the growth potential for their integration with PI applications. To date, very little attention has been paid to the role of EEG devices in the personal and lived informatics contexts described by I. Li et al. (2010) and Rooksby et al. (2014).

A survey study conducted by Hassib, Khamis, Schneegass, Shirazi, and Alt (2016) sought to investigate user’s needs for ‘bio-sensing’ across a range of wearable devices capable of capturing physiological, emotional and cognitive data. Participants were asked for their thoughts about acquiring and sharing data from various sources, including BCI headsets. They were found to be significantly more interested in collecting cognitive data (e.g. relating to stress, concentration and relaxation), than both physiological data (e.g. heart rate, blood pressure and body temperature) and emotional data (e.g. happiness, anger, sadness). The study also revealed that participants were interested in sharing and receiving cognitive and emotional data with (and from) others, primarily within their network of close friends and family.

Kunze, Iwamura, Kise, Uchida, and Omachi (2013) discuss the possibility of tracking people’s cognitive activity, in much the same way that services such as Fitbit are used for recording and analysing physical activity. Specifically, they focus on the use of eye-tracking for inferring cognitive activity during reading. The authors acknowledge EEG is a promising cognitive process measurement method. However, they favour eye-tracking for the reading task that their research is focused on, positing that participants’ eye-tracking data could be used as a measurement of ‘mental fitness’ by determining and tracking “how much people read, what type of document they’re reading, and how much they understand of what they’re reading” (Kunze et al., 2013, p. 106). The authors suggest that the idea of cognitive tracking can be applied to a wide variety of cognitive tasks to “enable new forms of self-reflection and suggest strategies to optimize mental fitness and well-being” (Kunze et al., 2013, p. 108).

Further research has also considered the possibility of including EEG data within a PI tool. Kido (2012) proposed, MyFinder, an application for recording a user’s mental state data from EEG, alongside other personal data such as events, comments, pictures and stress status. These studies, however, did not explore actual user experiences of consumer-grade BCI technology for acquiring and sharing such data, nor did they assess whether users’ expectations for the system, align with the actual capabilities of current consumer-grade BCI devices.

In Study 1 (Chapter 3), we seek to explore the capabilities that users would like to be

available in a PI system capable of recording EEG data. Following on from this we seek to address the challenges (C1-C3) found by our study in subsequent studies.

### 2.7.1 Fitbit for the Brain

There is a growing reference to using BCI technologies in the context of PI as ‘Fitbit for the brain’. One article published in the Irish Times (McCall, 2018) references work by BrainWaveBank<sup>19</sup> to build a device capable of diagnosing Alzheimer’s disease earlier as “Fitbit for the brain”. Similarly, work carried out at the University of Pennsylvania and published in PennToday describes their attempts to make a BCI device as being “akin to a Fitbit for the brain” (Berger, 2019). Additionally, published works Wilson et al. (2018) describes the design of a Cognitive Activity Tracker, as designing a “Fitbit for the brain”. The research conducted by Kunze et al. (2013) described earlier in this section, although not specifically using BCI technologies, also described their work using this metaphor.

This metaphor highlights how these technologies are already being considered by researchers, even if they are not being adopted widely by the public. Therefore, further research into how people might consider these technologies, when likened to a Fitbit, might be used for self-reflection and self-understanding may be beneficial.

### 2.7.2 Commercially available technologies

Muse and Emotiv both have apps available to download for consumers which enable tracking of data recorded by their headsets. The Emotiv application, myEmotiv<sup>20</sup>, enables users to capture data relating to “6 cognitive metrics” (Emotiv, n.d.):

- Interest - Measures how much you like or dislike something
- Excitement - Measures your level of mental arousal
- Relaxation - Is your ability to switch off and reach a calm mental state
- Engagement - Measures how immersed you are in what you are doing or experiencing
- Stress - Measures how comfortable you are with the current challenge you are facing
- Focus - Is your ability to concentrate on one task and ignore distractions

---

<sup>19</sup><https://www.brainwavebank.com/>

<sup>20</sup><https://www.emotiv.com/myemotiv/>

Muse allows the tracking of meditation sessions recorded using a combination of their app and headset and provides users with feedback about how much time was spent in either active, neutral and calm meditative states. Depending on the version of the Muse headband used, users also have access to additional data such as heart rate, breathing rate, and posture provided by a variety of other sensors built into the headband.

These technologies are probably still confined to those who are eager to try new technologies rather than being adopted by large portions of the population. However, there is limited evidence available from these companies about how they have developed these technologies and how people that are using them to understand themselves better do so.

### **2.7.3 Neurotechnology & Neuroergonomics**

Technologies that use BCIs as part of their setup are now becoming known as neurotechnologies and a related field of study that may have implications for BCI PI systems is ‘neuroergonomics’ which applies research from the field of neuroscience to ergonomics. Research from the field of neuroergonomics could provide insights for the field of PI in relation to using neurotechnologies for self-tracking performance and related indicators.

In considering neuroergonomics in relation to human physical performance, Rahman, Karwowski, Fafrowicz, and Hancock (2019) conducted a systematic review of applications of electroencephalography to physical activities. Their research highlights an extensive collection of activities within which researchers have investigated the use of EEG. Their work also highlights some research that incorporates the use of neural correlates in combination with other physiological signals such as electromyography and electrocardiography, for measuring electrical activity in muscle and the heart respectively. Topics of research uncovered by the researchers include “Muscular Activities with less mobility”, “Physical activities With Cognition” and “Miscellaneous Physical Activities”. The first topic deals with using EEG data in relation to activities that required limited or restricted movements such as just single joints or limbs. The second topic deals with EEG research related to tasks that required a combination of physical and cognitive demands. Finally, the researchers highlight a variety of EEG research studies that involved activities such as yoga, walking, running, swimming. The authors suggest that review demonstrates:

the substantial role of brain activities in controlling performance, fatigue, preference, emotion, cognition, and perception in relation to physical move-

ment. EEG signatures also have significant correlation with load levels, intensity, modes, stages of task preparation and execution (Rahman et al., 2019).

This area of research highlights the growing trend to consider EEG data as viable source of information. It can be considered in relation to PI in much the same way that people who use PI systems to track their activities to become more fit. Neuroergonomics could be considered as an extension of things that people may wish to track.

## 2.8 Trust in PI

Research related to trust in technology has been conducted for many years. Often, this research focuses on trust in the automation of a process, whereby a technology is supplementing or augmenting a process with potential implications for the user. An example where trust is of particular importance perhaps is critical systems whereby someone must trust that the technology is accurately providing information so they can take appropriate actions (Cioroiaica, Buhnova, Kuhn, & Schneider, 2020).

Some of the earliest work concerning developing an understanding of trust between humans and machines is a model developed by (Muir, 1987), their model is defined as follows:

Trust is the expectation, held by a member of a system, of persistence of the natural and moral social orders, and of technically competent performance, and of fiduciary responsibility, from a member of the system, and is related to, but not necessarily isomorphic with, objective measures of these qualities. (Muir, 1987)

Muir’s work also highlights several key points of interest drawn from previous research, such as people’s increase in trust is found to be related to “[their] ability to estimate the predictability of the machine’s behaviours”. Based on peoples’ tendency to overestimate representativeness of small samples, Muir also suggests that they determine levels of predictability on the basis of unreliably small sample sizes. Furthermore, as trust in a system develops, this will likely reduce the amount of sampling performed by a person to further determine and calibrate that trust. Paradoxically, Muir suggests therefore that a person’s knowledge about a machine is likely to be inversely related to their trust in the machine. Another factor that Muir suggests will influence people’s trust is based on the machine’s degrees of freedom, suggesting that the greater the degrees of freedom the less likely people will be to trust it. Two other factors that Muir argues

may that affect trust, are the environment that the machine works in, and the ability of a human to observe that machines behaviour (Muir, 1987).

The concept of trustworthiness is described by O'Hara (2012), as being the characterisation that, all things being equal, an agent (someone or something), will do what it claims it will do. O'Hara describes trust as the consideration by a person of the trustworthiness of an agent. Trust is considered to be an attitude towards the trustworthiness of another (O'Hara, 2012).

In the context of Personal Informatics, it may be considered that sensing technologies are relieving operator burden of having to manually track aspects of their lives. With the pervasive role that technologies play in people's lives and their increasing reliance on technologies, the ability to trust the technologies that we use plays an important role. Although a self-tracker may not be operating a critical system, incorrect decisions made based on technology-provided information could perhaps have unintended consequences. In particular, when people are using technologies to record data about themselves (e.g. health related) that they may then use to infer things about themselves or take actions based on these results, their ability to trust they are being provided with accurate information is of utmost importance.

Previously, researchers have found that when using automated systems people were more trusting, in that they were less likely to verify the system state when a system displayed its confidence in being correct about its analysis. The higher the displayed confidence (e.g. values greater than 70%) resulted in participants seeking to verify the automated result less often, whereas values with low or medium confidence resulted in participants wishing to verifying the results more (Antifakos, Kern, Schiele, & Schwaninger, 2005). In the context of Personal Informatics, similar system confidence implementations as these can be seen in Exist.io's inclusion of 'confidence ratings', appearing as star ratings, based on the statistical significance of the correlations presented to users.

Rupp, Michaelis, McConnell, and Smither (2016) sought to understand the trust that people had in fitness trackers. The authors conducted a lab study whereby participants were asked to interact with one of six types of fitness trackers, in a manner which a person would in everyday life. Afterwards, participants were asked to complete three scaled questionnaires: the Trust in Automation scale (Jian, Bisantz, & Drury, 2000), the Wearable Technology Trust scale (Rupp et al., 2016), and the Wearable Technology Motivation scale (Rupp et al., 2016). The authors found that there was a direct impact of trust and intrinsic motivation, suggesting that people's trust in a device has an

impact on the person’s motivation to continue using the device.

To assess whether people trusted the accuracy of pedometers, Huppert, Kranz, and Hoelzl (2019) conducted an in-the-wild study where participants were asked to wear a wrist-worn pedometer, for which the researchers were able to alter the presented step counts. The authors found that none of the participants noticed that the values they were presented with had been modified so that they were increased or decreased by 15% compared to a baseline measurement. This suggests that people may have an implicit trust that the values produced by a provided fitness tracking technology are correct, or are unable to tell specifically when the values presented are inaccurate and should be questioned.

Therefore, it is unclear what impacts the level of trust that people place in wearable sensing devices. Obviously, if a device displays zero and the wearer knows they have taken some steps then one might assume that person is unlikely to trust the device. However, as suggested by Huppert et al. (2019), the threshold at which someone will begin to not trust a device may be significantly outside the range of what they perceive to be a valid value.

Before using a device or technology, people must often select a device and consciously decide to engage in its use. One factor that influences a person’s decision to choose a device, may how the device has been advertised. This may result in someone visiting the manufacturer’s website to review a product and gain a better understanding of what that product can do, before deciding to purchase. However, *commercial* tracking technologies used for PI do not always endure the same level of scrutiny as medical devices used for obtaining similar kinds of data. Therefore the veracity of the claims that commercial device manufacturers make, or information they may omit, is perhaps important in determining whether to trust the device. Coates McCall, Lau, Minielly, and Illes (2019) looked at the ethical implications of claims associated with commercially available wearable brain technologies. The researchers found that only 8 out of 41 devices had links to peer-reviewed papers that backed up the claims made with evidence. They advocate for the important role that scientific evidence plays in substantiating claims and informing public trust.

This research highlights current and existing PI systems but does not provide insights into how those who are new to these technologies determine or decide how these devices are trustworthy. We seek to further understand this with our lab-based study presented in Study 4 (Chapter 7), as well as specific interview questions in Study 2 (Chapter 4) and Study 3a (Chapter 5).



## 2.9 Chapter Summary

In this chapter we reviewed the current focuses of Personal Informatics research. We provided background insights into the precursors to PI, such as the increased ubiquity of sensing technologies. We provided an overview of the main streams of research related to understanding those who self-track and the ways in which users may track data to better understand themselves. We highlight work by I. Li et al. (2010) and Epstein et al. (2015) which provides models for the way people track and attempt to make sense of data within their PI-tracking practices, providing the ‘*Stage-Based*’ and ‘*Lived Informatics*’ models of Personal Informatics, respectively. The other streams of research within PI that we explored are those of recommendations for design features and visualisation techniques. To complete the background information on PI we described ‘Personal Informatics systems’, with examples of currently available tools, such as those provided by Fitbit, Google Health and Apple.

To position the research described in this thesis we defined ‘novel sensing technologies’. We highlighted the increasing availability, decreasing cost and wearability of these sensors. We also highlighted the nature of data collection in relation to PI as automated, unobtrusive, multifaceted and personal in both its collection and its analysis. We also discussed that the data produced by these sensing technologies may have varying levels of trustworthiness, transparency and that users may lack familiarity with the data being collected.

Additionally, we provided background information in relation to the sensing technologies used in our research, and our reason for choosing a BCI as an appropriate technology probe. We also looked into the current use of BCIs within the domain of Personal Informatics, references to BCI technologies as ‘Fitbit for the Brain’, and the emerging fields of neurotechnology and neuroergonomics. Finally, we provided an overview of trust within the domain of PI. Firstly by describing background information in relation to trust in technology, which has been studied for many years and moving toward more recent studies related to trust in data, wearable pedometers and trust in neurotechnologies.

In the following chapters we will describe the results of the research conducted for this thesis centred around what people expect from a BCI-based Personal Informatics system (Chapter 3), the challenges of making sense of novel sensor technology data within the context of PI (Chapters 4 & 5), different ways of aiding exploration of PI data (Chapter 6), and the role data visualisation may play in the formation of trust in sensing technologies (Chapter 7).

## Chapter 3

# Fitbit for the Mind?: An Exploratory Study of Personal Informatics

*The following chapter is an unabridged version of an extended abstract paper that was accepted for publication at CHI 2018: (see Appendix A).*

*Dudley, C., & Jones, S. L. (2018, April). Fitbit for the Mind? An Exploratory Study of ‘Cognitive Personal Informatics’. In Extended Abstracts of the 2018 CHI Conference on Human Factors in Computing Systems (pp. 1-6).*

### 3.1 Chapter Overview

As discussed in the previous chapters, Personal Informatics (PI) and life-logging systems allow users to track data about their everyday activities and behaviours, and explore the collected data in order to uncover meaningful insights about themselves (I. Li et al., 2010). These systems often enable self-reflection and support individuals to change or improve aspects of their lives (Choe et al., 2014). Typically, PI systems are used to aggregate and analyse data relating to life-facets such as physical activity (from services such as Fitbit, Google Fit and Apple Watch), diet (e.g. from MyFitnessPal, LoseIt and Cron-O-Meter), productivity (e.g. from RescueTime), and mood (e.g. MoodTracker). However, the variety of sources from which data can be collected is continuously expanding due to the emergence and availability of new wearable sensor technologies.

Consumer-grade brain-computer interface (BCI) headsets, although currently in their infancy, may present an opportunity for the average consumer to track electroencephalogram (EEG) data, or ‘brain data’, offering users a figurative ‘Fitbit for the mind’. This raises a myriad of questions about the use of EEG data in a personal informatics context. What value do users believe they can gain from recording EEG data? What problems are people likely to experience when current consumer-grade BCI technologies are used for self-tracking? What HCI research challenges do we face in integrating BCI/EEG technologies with personal informatics systems?

In this chapter we seek to understand design opportunities, challenges, and technical, social and ethical implications for the “near-future” technology of ‘Cognitive Personal Informatics’ (CPI) — a class of tools that enables users to collect and analyse EEG data for the purpose of understanding and monitoring their brain activity. We report a study in which participants were provided with a consumer-grade BCI headset and a mobile application capable of recording and visualising their EEG data in real-time. Using qualitative data from interviews with participants following their use of the technology, we assess their initial reactions and investigate their expectations for the value and insights that future Cognitive Personal Informatics systems will provide. We

also identify scenarios in which participants anticipate this type of technology being used, possible barriers to use, and concerns that are expressed.

While the majority of participants were enthusiastic about their experience using the cognitive personal informatics system, and suggested a number of reasons for using such technology beyond the confines of our study, our participants’ feedback also highlights four significant challenges that must be overcome in order for the cognitive personal informatics systems to become valuable tools for mainstream users. These are: addressing the indirection between meaningful psychological states and brain activity data, supporting diverse tracking styles, encouraging exploratory and enquiring analysis approaches, and overcoming misconceptions about the brain. Our contribution is to provide an understanding of: (1) the motivations that are likely to drive EEG tracking and (2) the challenges that designers are likely face in providing usable personal informatics systems that integrate ‘brain data’.

## **3.2 Methodology**

We designed a study to elicit participants’ views on how and why cognitive personal informatics system might be used, and to capture their expectations for features that future systems may include. We designed a system to provide participants with an experience of tracking brain activity data whilst going about their normal activities, much like wearing a Fitbit or other consumer tracking device. Our system comprised an Android application called ‘Cognition Tracker’, which was developed in order to receive data from NeuroSky’s MindWave Mobile BCI device. The lightweight, wireless nature of the MindWave allows it to be used in movement and normal activity. Additionally, having fewer electrodes (and the electrodes being dry) makes it easier than the Emotiv EPOC for first-time users to handle or adjust by themselves. Ethical considerations for this study followed the Department of Computer Science’s 13-point checklist, see Appendix G. Data pertaining to this study is retained in accordance with the data management plan found in Appendix F.

### **3.2.1 Cognition Tracker App**

The Cognition Tracker application (Figure 3-1) provided a simple line graph visualisation of the five wave bands recorded by the headset; Alpha, Beta, Gamma, Delta, and Theta, plotted over a 60 second period and updated in real-time every second. In addition, two values representing meditation (mental calmness/relaxation) and attention (mental focus) derived from NeuroSky’s proprietary algorithms (NeuroSky, n.d.),

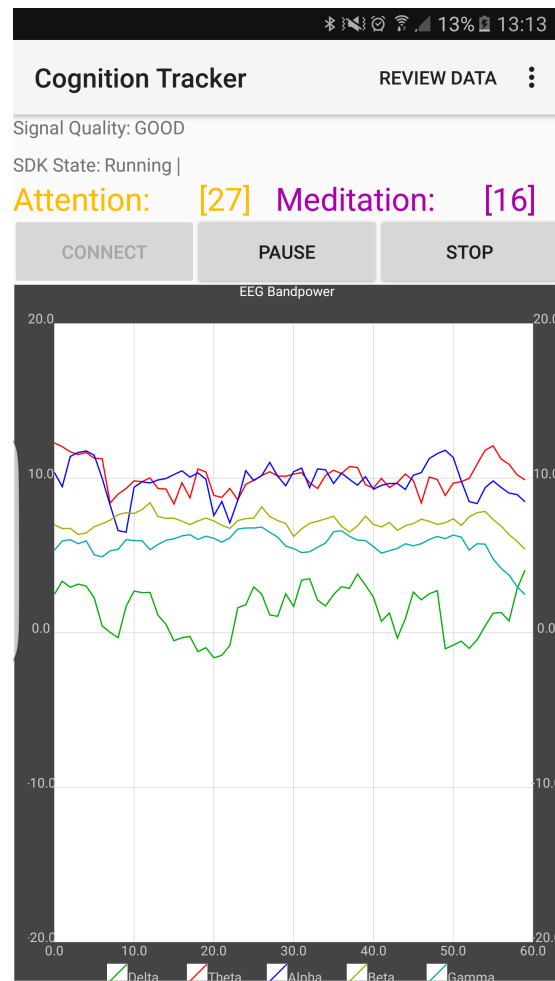


Figure 3-1: The Cognition Tracker Android application showing NeuroSky's eSense values (Attention 0-100, Meditation 0-100) and five EEG wave band powers (alpha, beta, gamma, delta, theta) in real-time.

were displayed above the graph and also updated in real-time. The application also contained a tabbed interface that allowed users to switch to a historical view of the data recorded, providing a time-stamped log of wave band, meditation and attention values. The data visualised represented 'raw data' from the NeuroSky headset and was not interpreted or transformed for the participants in any way by the software. The application served to give a practical demonstration of a BCI device acting as a real-time data tracker, rather than an input/control device. The application did not provide participants with a fully implemented PI system with extensive data exploration or analysis features; rather the application provided basic visualisations and a log of the data such that participants could appreciate the type of data that was being recorded and consider the potential uses or future iterations of such an application.

Participants were encouraged to consider this as a starting point for a future cognitive personal informatics technology and make suggestions about additional features or use cases.

### 3.2.2 Participants

16 participants (9 male, 7 female), aged 21-62 ( $M=30.08$ ,  $SD=10.69$ ), were recruited, via posting on the University of Bath’s online noticeboard and by word of mouth. No specific requirements were needed for participation in the study. Participants had wide-ranging previous experiences with PI systems. Seven participants (P1, P3, P7, P11, P13, P14, P15) had previously used fitness trackers. One of these participants (P1) mentioned having used a range of tracking devices, including wearable fitness trackers and online services for ‘life-logging’ over a two-month period. None of the participants were aware of existing technologies that could be used to track their EEG data for personal use. When asked about technologies that they were already aware of for personal data tracking, participants provided examples for heart rate trackers, eye trackers, smart watches and wearable activity trackers, but none that focus on cognitive data.

Participant #	Age	Gender	Occupation
P1	32	F	PhD Researcher
P1	32	F	PhD Researcher
P2	28	F	EngD Student
P3	24	F	PhD Student
P4	30	M	PhD Student
P5	26	M	Researcher
P6	21	M	PhD Student
P7	25	M	PhD Student
P8	24	M	PhD Student
P9	24	F	PhD Student
P10	28	F	PhD Student
P11	25	M	PhD Student
P12	23	M	PhD Student
P13	23	M	PhD Student / Professional Cyclist
P14	32	F	Physiotherapist
P15	62	F	Retired Retail Assistant
P16	48	M	Quality Assurance Specialist

Table 3.1: Study 1 - Participant Demographics

### 3.2.3 Procedure

Participants were provided with an information sheet and consent form (see Appendix B) to read and sign prior to starting the study. All participants were given a NeuroSky MindWave Mobile headset and a mobile device with the Cognition Tracker application installed to use for around 30 minutes, either at home or at work. 30 minutes was deemed an appropriate target based on pilot testing, representing a suitable period for participants to get a sense of the data being captured and the features of the application. It also mitigated the possible discomfort of the headset during prolonged use, and kept the typical duration of the participants' involvement in the study at approximately one-hour, including a follow-up interview. The study was run over a period of 14 weeks.

The participants were not given any specific instructions as to how they should use the system. Rather, participants were told they were free to use the application however they saw fit. Participants were given an initial introduction by the researcher, explaining how the application worked, what was presented on the display, and how to ensure the headset was transmitting correctly. Participants were then provided with help fitting the headset to ensure that it was positioned correctly, with a good quality signal connection, and that they knew how to begin recording data. Participants were then free to undertake any activity, e.g. going about their normal activities, whilst wearing the headset and having access to the Cognition Tracker application with the live data stream and historical data log. The researcher was not present at this time. Participants mentioned working, reading, listening to music, watching a movie, playing games on their phone, eating, and cooking as tasks that they carried out during their time with the system. Participants 14, 15 and 16 spent time use the system at home. P1, 2, 3 and 4 spent time using the technology outside (e.g. whilst going for a walk) during their free time, and all other participants used the system at work in an office setting. P1, 2, 6 and 13 all reported spending some time engaging socially with others during breaks from work, whilst wearing the headset and recording data. Participants took part in an interview shortly after using the headset and application. The interviewer asked participants about their initial experience with the system, and who they thought might use and benefit from a system for recording EEG data. Participants were also asked about their interactions with the Cognition Tracker application, and if they had noticed anything interesting or intriguing in their data.

Participants were prompted to identify any questions or hypotheses that they felt their EEG data might enable them to answer, and if there were other types of data they would consider combining with EEG data to learn more about themselves. Participants

Category of Use	No. of Participants
Improving self-understanding	8
Monitoring medical conditions	8
Optimising behaviour/performance	6
Hobbyist/technophile uses	5
Supplementing existing tracking technologies	4
Monitoring general health and well-being	3
Improving understanding of others	1

Table 3.2: Potential categories of use for CPI systems

were asked to consider how they might use this type of technology in the future and if they had any concerns about recording their EEG data. All interview audio was transcribed and then coded and thematically analysed (Braun & Clarke, 2006). Initial codes were generated by the researchers reviewing the transcribed interview data in full. Once initial codes were generated using an inductive, reflexive approach, these were filtered and combined to generate overarching themes.

### 3.3 Interview Results

#### 3.3.1 Why Use Cognitive Personal Informatics Systems?

During the interviews, participants were asked to consider and explain the reasons why they would use cognitive personal informatics systems that incorporate EEG data. The primary purpose of this was to discover the meanings that people ascribe to the data and explore anticipated uses of the data, rather than to assess the validity of suggestions by participants. Thematic analysis of their responses revealed seven high-level categories of use. These related to: health monitoring for specific medical conditions; monitoring general well-being; improving self-understanding; optimising behaviour/performance; improving understanding of others; supplementing existing tracking tools; and hobbyist/technophile uses.

##### 3.3.1.1 CPI for Monitoring Medical Conditions

Half of our participants suggested reasons for using cognitive personal informatics that centred on medical diagnosis and condition management. Many saw the system as a potential tool for supporting the identification and treatment of mental health conditions such as depression and anxiety, or factors such as stress that could allow mental health conditions to develop. For example, P4 suggested that “if people are really in stress or going through depression, you could maybe try to ‘backtrack’ in some sense”.



Some participants believed that the tool might be capable of detecting or presenting early signs of brain disorders such as dementia, for example P15 who has a family history of dementia stated “to avoid getting dementia, knowing what makes your brain work and what keeps it active would be a useful tool”. P9 suggested that CPI would benefit “people who care about their health or have a history of [brain related] disease, and they want to monitor that.” P4 thought that it may be useful as part of a self-diagnosis tool for the elderly in particular “to check if their brains are still working in some sense, so it’s like a feedback loop to check the healthiness of people.” P2 suggested that health professionals could utilise the data from patients’ cognitive tracking tools to “determine whether the patients are being mentally active”. P16 suggested that since tracking physical activity, e.g. via Fitbit, “is all about improving physical health, I assume the practical benefit of this would be to improve mental health. That could be really valuable for a lot of people”. These responses highlight a willingness from participants to engage with BCI devices for mental health monitoring.

### **3.3.1.2 CPI for Monitoring General Well-being**

P15 had an interest in being able to record her EEG data to improve her general well-being, for example by improving sleep: “I don’t sleep that well and if I could use this to find a pattern where I become more relaxed and I knew that doing certain things, like say listening to music would work for me, then that’d be really useful”. P3 suggested that they may use it if they felt the need to generally improve their affective state: “It’s interesting but I wouldn’t like to use it every day. I can use it when I feel very down, in that specific time of my life.” P1 envisaged a relative using the device to determine causes of high blood pressure: “my mom has high blood pressure and she was wondering what it is that’s causing it, it would be quite nice [to find the cause].”

### **3.3.1.3 Improving Understanding of Self & Optimising Behaviour**

Several participants suggested that they would be likely to use cognitive data generated from an EEG device for self-reflection and improving their understanding of self, rather than addressing specific health or well-being issues. P2 suggested that cognitive personal informatics tools could be used as a “performance enhancement tool...to find out more about yourself in general, but also to see how you could use that information to get better”, as well as “to try to get the best out yourself, out of your mind.”

P1 stated that they would use such a system “...to get better at quick thinking, generating ideas quickly, being creative, spontaneous.” P3 reported that they would like the system to determine “my working efficiency today, and maybe it can tell me

the truth of - am I working or not?” and P11 was interested in recording BCI data “to gain a better awareness of how my brain or how my body reacts to certain activities.” and understand “in what situations do I perform best?”

One participant (P4) suggested being able to use EEG to better understand their personal tastes or preferences; using the EEG data to infer emotional responses to events/activities/objects. They also suggested that such data might allow machines and artificial intelligence systems to “gain a better understanding” of the user by automatically detecting their “likes and dislikes” and using this data to adapt and personalise the system.

#### **3.3.1.4 Improving Understanding of Others**

P11 suggested that openly sharing the data from cognitive personal informatics systems could “expose data so that you can understand the [mental] state of someone [else]”, and suggested that this would be particularly useful in cases where people found it difficult to communicate such information (e.g. “children” or “individuals with autism spectrum disorder” (P11)). P11 described transmitting the data to others as adding a “transparency layer for mental state”. Methods for sharing data from CPI systems may build upon previous work on visual communication of physiological data (e.g (Ashford, 2014)).

#### **3.3.1.5 Supplementing Existing Tracking Tools**

One of our participants described the EEG tracking device as a mental counterpart to existing physical activity tracking devices: “Fitbit could cover the physical side, and this could cover the mental side, that might be all you need [for holistic tracking]” (P15). Often, the EEG data was viewed as an additional source for improving the accuracy of other tracking technologies. For example, P5 and P7 described the possibility of using the device for better detection of physiological activity (e.g. detecting exercise, measuring heart rate, detecting sleep), supplementing wearable devices such as Fitbit – “I guess you could combine it with standard pedometers and stuff like that for better accuracy” (P5). P16 felt that it could be used in combination with “a number of sensors that detect when things start to go wrong with your body.”

#### **3.3.1.6 Hobbyist/Technophile Uses & Hedonic Experience**

Many of the participants were enthusiastic about the technology, reporting a sense of enjoyment from their experience using it. For example, P2 said, “I really enjoyed it, I felt like a futuristic sci-fi Robocop!” and “I find it fascinating to see what my brain is

doing.” P9 said “I think it’s quite cool to use to know a bit more about what’s going on inside.” P1 expressed positivity about the value that it could provide in future, despite some reservations about whether such value could be fully realised: “I don’t know if you could measure stress levels? Because it would be just amazing if you could do that.” P8 and P11 both described the technology as “important” in terms of the value that it could provide to users lives. P10 said it seemed “cool” to be able “to collect quite a lot of data and to monitor yourself.” Several participants made comments that were implicit of the ‘fetishized tracking’ reported by Rooksby et al. (Rooksby et al., 2014), whereby users are keen to engage with the system and evaluate the data that is generated because of hedonic experiences or general enthusiasm for technology, and because there is the availability of the device and tools to do so. P9 stated, “for now maybe more as a personal hobbyist you would just record new things and you would have some data and you can geek out a bit more.”

### **3.3.2 What Insights Will Cognitive Personal Informatics Systems Provide?**

Although participants were presented with two real-time metrics for meditation and attention levels, according to NeuroSky’s proprietary eSense algorithms, and a visual representation of five wave powers, many made assumptions about additional capabilities of the device for detecting other cognitive processes, states and measures from the EEG data, and suggested that they could be provided as numerical or categorical information for the user.

Table 3.3 summarises a list of the metrics, cognitive processes and psychological states, which participants envisioned being able to monitor with the use of a personal informatics system. Some metrics show participants considering the devices as ‘counters’, i.e. solely producing quantitative data, in the same way that fitness devices are step and calorie counters. Cognitive tracking devices were viewed as quantitative ‘stress counters’, ‘cognitive load counters’, ‘brain activation counters’, and so on. The participant’s suggestions demonstrate their expectations that there is a broad range of meaningful, quantifiable values that can be obtained from a CPI system.

Other suggestions implied that participants believed there was rich, complex, qualitative data that could be captured. For example, suggesting that such a device could provide insights about “what my brain is doing” or “what I’m thinking” (P2), their “mental state” (P11), or “what’s actually going on when people are trying to be creative” (P9).

Metric / Process / State	No. of Participants
Sleep rhythm	6
Focus / attention	5
Stress / relaxation level	5
What my brain is doing / mental state	5
Concentration level	3
Disease progression	3
Productivity / efficiency	3
Cognitive load/ mental strain	2
Current mood (e.g. anger)	2
Biorhythm	1
Brain activation	1
Consciousness / fainting	1
Creativity	1
Depression	1
Liking / preference	1
Meditation state	1
Mind efficiency	1
Praying state	1
Procrastination	1

Table 3.3: Suggested metrics/states for tracking using CPI

### 3.3.2.1 Analysing EEG Data in a CPI System

Participants suggested a number of different types of analyses that they would like to perform (or for the system to perform automatically) on the basis of the data that the system was collecting. These analyses often implied that data would be: recorded frequently, in a wide variety of circumstances, over long periods of time, and fused with other forms of data to provide meaningful insights.

The most common type of analysis suggested by participants involved the comparison of brain wave data across different activities, to determine the effects of each activity on the user’s cognitive state. For example, P1 was interested in seeing if different activities lead to different patterns in their EEG: “Maybe doing sports, then reading a book then maybe have a call with a relative...I expect this will lead to different patterns in the EEG... It would be really interesting to see how your brain behaves in certain situations.” Similarly, P7 was interested to see how his brain would respond to different activities: “it would be more out of curiosity just to see what happens to my brain when I do different things” and P14 wanted to find out about the effects of her environment on her mental state: “If I was to wear it for a longer period of time and maybe with like, in different environments, home environment, work environment,

social, I could kinda see where I'm most comfortable maybe, most relaxed".

Participants expressed interest in performing both inter-session and intra-session analyses. Inter-session analyses comprise comparing EEG data across distinct recording sessions, either for the same activity being performed at different times or in different settings, or comparing the data across recording sessions for different activities altogether. Intra-session analyses comprise a finer granularity of data being inspected in detail, for example drilling down into particular fluctuations in the EEG data within a single recording session and correlating them with particular external events to understand what effect they have on brain activity. P13 wanted to be able to switch between macro and micro level analyses, "zooming in" on interesting specific points within the data, e.g. a spike in attention values, and "zooming out" to see larger trends.

There was also interest from participants in being able to compare their own EEG data 'to the norm', e.g. "you could compare this EEG data with lots of other EEG datasets...if my EEG data is comparable to the average healthy participant I would think OK, my brain, or the way I'm thinking, seems to be fine" (P4).

Participant 2 was keen to analyse data to uncover temporal patterns that reflected perceived variations in cognitive function, which they referred to as 'biorhythms'; "[You could use this] to notice your biorhythm, to see when you're most productive, to try to get the best out yourself, out of your mind..."

Four participants indicated that they would like to obtain summaries of their cognitive activity in the form of high-level information to accompany low-level EEG data, for example showing "summary statistics for each individual wave" (P13), daily values such as "maximum time spent concentrating and average amount of time spent concentrating" (P3), or statements such as "you have been very focused today" (P7), and "your EEG contains early warning signs that you might need to see a doctor" (P16)).

Analysis for the identification of triggers; external stimuli that activated a certain response in brain activity, was of importance to several participants, e.g. "I could potentially, by identifying that I find some things more stressful, find ways to try and limit that, to some extent" (P14).

Several participants reported experimenting with the Cognition Tracker tool during the study, deliberately altering their behaviour and observing the output in order to try and understand how changes in their actions were manifested within the data. E.g. "I felt like I could separate certain waves by doing certain things...I wanted to just see how what I do has an effect on these values" (P13), "It felt like I could control the

brain wave chart just by altering the way I was thinking, it seemed to correspond with something that was going on” (P16).

Often the identification of triggers implied the need for extremely rich contextual data collection, alongside EEG tracking. Participants wanted to be able to identify notable events within their EEG data and study the relationship with data that revealed contextual information about the event. For example, P4 suggested combining EEG with eye-tracking data: “...you could connect every visual impression with your biophysics and your brain activity, that would be interesting.”

### **3.3.2.2 Data Integration**

Several existing PI systems allow for multifaceted data collection and analysis. For example, Exist.io provides users with correlational insights such as “on days where listen to more music, you walk more” by aggregating multiple personal data streams (e.g. Fitbit steps and Last.FM song plays) (S. L. Jones & Kelly, 2018). Participants were asked to consider whether they would integrate other types of data with their BCI data, and explain how this might provide additional insights. The types of data discussed, and reasons for combining them, are shown in Table 3.4

### **3.3.3 Barriers to Use for Cognitive Personal Informatics**

Based on their initial experience with the system, participants explicitly mentioned several potential barriers to use and concerns about the technology (see Table 3.5). Some of these barriers and concerns were experienced first-hand during the study, whilst others were predicted to be salient for future usage.

#### **3.3.3.1 Ergonomic & Aesthetic Hardware Issues**

During the study, appearance and discomfort of the headset was a common concern shown by multiple participants. P5 said that the NeuroSky headset was “uncomfortable on the ear lobe” and participants 1, 4, 10, 11, 13, 14 and 15 all reported minor discomfort either on the forehead or ear.

Four participants also expressed concern that the NeuroSky headset was not suitable to wear in public, with P6 stating: “I looked in the mirror and it looks funny.” P3 highlighted that they would not be inclined to wear the headset in front of others who may not understand the technology, “such as the elderly”, saying they would not use the headset “in front of my grandparents because it might scare them.”. P1 highlighted that there may be cultural factors that would make the use of BCI devices less likely

Data	Reason
Blood Pressure	It was suggested that combining blood pressure data with EEG data may support the evaluation of stress e.g. “my mom has high blood pressure and she was wondering what’s causing it, it would be quite nice” (P1)
Heart Rate	Heart rate data was suggested for comparing physiological with cognitive responses, “I felt my heart leaping a bit with joy and then the brain activity didn’t seem to change much” (P2), or provide additional insights to “track the chances of having a stroke” (P12)
Galvanic Skin Response	Combining with EEG data to support evaluation of stress levels (P1)
Browser History	Web browsing data could be correlated with EEG data and used for marketing, surveillance and monitoring “you could find out some interests and favours like people process certain input like media or movies”, “if your boss knows that you have certain favours or certain tastes... it might not come in handy” (P4)
Task/Activity Type	Participants suggested that recording activities (e.g. tasks and behaviours) alongside EEG data, could enable understanding of how brain activity varies in different circumstances, “when you go do fitness or something you might want to check how your body reacted during the activity I would imagine.” (P7), as well as help to show “how [cognitive data] manifests itself in behaviour” (P9)
Physical Activity	Compare physical and mental “stress” (P15), Recording physical data could allow people “to see if physically your body is stressed but your mind could be effortless, or vice versa.” (P14)
Diary	Diary entries could provide further context to EEG data, and vice versa, “people write diaries about what they do, and how they feel” (P14). Participants felt that EEG data could enhance diary documents.
Gaze/Webcam/Video	Eye tracking or visual information could provide context to cognitive data by noting “what you’re paying visual attention to” (P13)

Table 3.4: Participant suggested additional data types for combining with EEG.

Barrier to Use	No. of Participants
Ergonomic and aesthetic issues	14
Difficulties Interpreting EEG Data	5
Privacy/Security Concerns	5
Health Concerns	3
Interference of Real-Time Feedback	2

Table 3.5: Barriers to use and concerns about CPI systems

- "...if you go to my home country people will look at you and be like, what is wrong with you?". P4 reported "people are reacting kind of strange when they see you having this on your head".

Generally speaking, current consumer-grade BCI devices are not discreet. They are large and obvious to others. Fitbits and other smart-band devices, although obviously smart-technology when looked at closely, tend to blend in as either wristbands or watches allowing for the discreet recording of data. Further research is required to develop discreet means of recording EEG data. One example of a system that may provide a discreet method for recording EEG is the Ear-EEG by Looney et al (Looney, Kidmose, & Mandic, 2014).

### 3.3.3.2 Difficulties Interpreting EEG Data

There were significant issues with data interpretation, P4 stated "you see basically in real-time how the EEG is captured which gives a good indication that something is going on, but I can't interpret those lines", referring to the wave band values in the graph. Participants were often unable to associate a particular cognitive process or mental state to changes observed in the wave band values. For example, "I was writing on the whiteboard all the waves certainly went down for some reason, I'm not sure why" (P2). P12 did not understand what the graph of EEG wave bands meant, "the graph was cool but I don't know what the lines meant" and P7 felt that exposing underlying wave bands within the visualisation was too technical: "it seems that it is aimed more to a technical person...". P7 suggested that a "normal" user "would like to receive feedback that...analyses it [the EEG data] and says - 'you have been very focused today'". Very few of our users had existing knowledge about EEG or the connotations of different wave bands. It is interesting to note self-experimentation was carried out by several participants, for example P13 said, "I felt like I could separate certain waves by doing certain things, but I'm not really sure...I couldn't really see any trends." In general participants found that the EEG wave powers did not provide an intuitive signal that could easily be associated with particular actions, thoughts or



mental states.

Separate from the visualization of the data, one participant, P4, mentioned being unsure about how accurate and reliable the data collected from the device would be. They said they “would really want to know how reliable the measured data is to making prediction or to making some statements about my state mentally or psychologically.”. Another participant, P15, specifically stated, however, that they “had no reason to doubt that it was accurate.”

### **3.3.3.3 Privacy and Security Concerns**

Privacy and surveillance were of concern to some participants. P5 stated they had the “usual concerns of who is using this data and for what [purpose]”. P11 echoed this saying that although they “find it quite cool” they would be concerned “about privacy, like who has access to the recordings”. Other participants did not show the same concern. P15 said, “Sharing data wouldn’t bother me. I don’t see any issue with it or anything secretive about it”. P9, although acknowledging that there may be privacy concerns, accepted this as a trade-off for using this type of technology: “People might find that the whole data privacy thing [is an issue]...but I think it’s quite cool to use to know a bit more about what’s going on inside”.

P4 expressed concern that other parties would have an interest in using the data for surveillance “from the boss who employs you to the health insurance which insures you...”, “It makes you more non-opaque. The more information you expose to others the more vulnerable you get, maybe...so it’s like a dangerous thing, what can you extract or infer from the information given?”

### **3.3.3.4 Health Concerns**

Three participants mentioned being unsure about the health risks from using BCI devices, despite non-invasive EEG recording having no known risks associated with it (NHS Choices, n.d.). P2 was concerned about the device transmitting signals into her brain, “I’m worried that the electrical signals might affect my brain or my body in general”. P8 said, “at some point I had some tickling in my ear, I don’t know if that’s normal.” P12 was concerned about potential physical dependence asking, “Is it healthy? I don’t know...people might become dependent on it.” Participants had been advised during their initial briefing that the experiment was considered ‘minimal risk’ and that they were free to remove the device and end the study at any time should they experience any discomfort or concerns.

### 3.3.3.5 Interference of Real-Time Feedback

In terms of monitoring themselves using the application it was noted by participants that attention to the mobile application was required to evaluate the EEG. This was described as distracting and presented a problem in determining attention and meditation scores. Often an attentive or meditative state was lost by having to concentrate on the device rather than the task at hand which the participants wanted to evaluate. P4 suggested not to “have too much interaction with the app...you should actually minimise the handling with the app.” and P13 felt that “to some extent if it can tell you how focused you are. I think that’s not going to help you because then you’re distracting yourself by checking how focused you are.”

P13 stated “I was very wary, for example when I was doing meditation I was like, OK I’m meditating and then I’d be like OK check the graph quickly before it changes and then suddenly it just spikes up again... So yeah I’d want to be able to scroll back [through the data at a later time] and see. I think general graph, analytic-type user interface stuff would just be handy for that sort of thing”. Real-time evaluation would likely benefit from the implementation of glanceable, or heads-up, technologies that do not distract users from what they are doing, but provide some insights with immediacy. A real-time feedback system would also likely require an algorithm capable of removing ‘noise’ generated by a user’s attention being drawn to the display.

Not all participants attended to the Cognition Tracker smartphone application during their time with it. P7 stated that they “didn’t really use it...I was just checking the signal and that’s it...”, instead he opted to collect data with the expectation that it could be reviewed or analysed later, for example, in the form of daily summaries. P4 said, “It would be cool just to capture while your app was running and [when] something was realised [the app] just says ‘Could you just mention what happened?’”. This suggestion implies that the onus of identifying notable or interesting insights within the data be transferred from the user to the machine. P4 thought that this might support more efficient contextual data collection, since the user could be prompted to provide additional contextual data only if the machine detected portions of the EEG data that were worth contextualising.

## 3.4 Discussion

The following sections draw upon our interview findings to identify and discuss four challenges that designers of cognitive personal informatics systems are likely to face.

### 3.4.1 Challenge 1: Addressing the indirection between meaningful psychological states and brain activity data

Our results highlight an apparent gap between users' expectations for cognitive personal informatics systems and current capabilities. Many of the metrics, psychological states and cognitive processes which participants envisioned tracking (see Table 3.3) currently lack published scientific studies demonstrating their accurate inference using devices such as the MindWave. Our study demonstrates that users are unable to reliably infer meaning from observations of raw EEG data by inwardly correlating observed changes in wave powers with particular activities or mental states. EEG data must therefore be processed, analysed and presented according to a superstructure of meaning (a mapping between EEG data and meaningful outputs) in order to provide understandable, actionable insights. Our study identifies a range of measures such as creativity, productivity, affective state, focus, and relaxation/stress that further research could seek to incorporate into such a mapping. Previous work has demonstrated feasibility for detecting some similar measures using medical grade EEG devices (e.g. creativity (Shemyakina, Nagornova, & Ponomarev, 2010), cognitive and memory performance (Klimesch, 1999), cognitive preparedness (Angelakis, Lubar, Stathopoulou, & Kounios, 2004)), however, further research would be required to evaluate these metrics in relation to consumer-grade BCI devices. Advances in the application of machine learning techniques for determining users' state from EEG data, as demonstrated in Chee-Keong Alfred and Chong Chia (2015), may help to gain traction in this area, along with further research focused on the design of interfaces for data exploration visualisation and analysis, enabling users to identify associations between patterns in EEG data, meaningful mental states, and external stimuli.

In order for users to gain reliable insights from a CPI system they must be provided with data that is accurate and valid. Only one of our participants expressed concerns about the accuracy of the data being presented to them. The displayed eSense values for 'attention' and 'meditation' exemplify black-box mappings between brain activity data and meaningful psychological states. In line with previous work (Lawson, Kirman, Linehan, Feltwell, & Hopkins, 2015), we found that users were "generally unconcerned about the scientific basis" upon which the technology provided information, and did not call in to question the reliability of its methods for converting raw signals to meaningful metrics. The trust placed in system outputs by users highlights the importance of ensuring that the system does not risk misinforming users.

### 3.4.2 Challenge 2: Supporting diverse tracking styles

Previous research has suggested that personal informatics systems are used in a number of overlapping styles (Rooksby et al., 2014). Our results suggest that cognitive self-tracking systems are also likely to be used in diverse ways. Participants formed expectations about potential uses for cognitive tracking tools based on their experiences with other tracking technologies (e.g. physical activity trackers). For example, the idea of directive tracking (Rooksby et al., 2014), driven by achieving a particular goal, e.g. 10,000 steps for physical activity, was considered in the context of EEG data. P16 asked: “what would be the goal... A certain number of thoughts? A certain amount of time concentrating? What is it counting and what are the recommended targets for those things?”

The framing of CPI systems as tools for directive tracking raises questions about the specific goals that they might support. While there are well-known daily guidelines for other aspects of human behaviour (e.g. steps for physical activity, kcals for food intake, hours for sleep) there exist few similar guidelines for the types of activity that our participants anticipated the CPI system being capable of tracking (e.g. ‘concentration’ or ‘focus’ levels). As a future research direction, our findings lead us to suggest that the design of cognitive personal informatics systems which enable directive tracking should be aligned with mental health and well-being recommendations where appropriate. However, in the likely absence of advised absolute values for certain aspects of mental activity, systems could consider either deriving norms from the analysis of data provided by other users, or orient the user towards relative goals and targets (e.g. “spend more time concentrating this week than you did last week” or “achieve your longest period of relaxation”) on the basis of user profiles. Previous research has demonstrated that consumer-grade EEG devices can be used to detect variations from an individual’s norm over time in specific situations (e.g. measuring stress in Crowley et al. (2010)). We argue that similar approaches should be incorporated into directive tracking features for CPI systems.

In addition to directive, goal driven tracking, our findings revealed that some participants were interested in using CPI systems for documentary purposes (e.g. augmenting diaries for reminiscing or recalling events) and diagnostic purposes (tracking to identify a specific problem, and uncover its root cause). In both cases, participants suggested using EEG data in combination with diverse forms of additional data, both to add context to their documented activities and to capture potential contributing factors for the problem being diagnosed. The need to provide rich contextual data in order to augment EEG data presents a significant challenge for future research.

### 3.4.3 Challenge 3: Encouraging exploratory & enquiring approaches

Our results showed that most participants expected to gain a deeper understanding of themselves by recording their EEG data within a CPI system. Users were often keen to explore and analyse their data at micro-levels (zooming in on particular moments) and macro-levels (gaining overviews of entire tracking sessions) in order to uncover meaningful insights. Users are therefore likely to benefit from visualisations and analytical tools that provide high-level overviews and the ability to drilldown into their data to see as fine-grained detail as they need. To date, few studies have evaluated existing visualisations for analysing and exploring brain data in this way. However, there were also expectations that the system would be capable of providing definitive diagnoses and binary outputs (for example classifying an EEG trace as either ‘good’ or ‘bad’, ‘healthy’ or ‘unhealthy’, and alerting the user accordingly) without the need for data exploration. While EEG is used clinically in the evaluation of dementias and encephalopathies, the limitations of consumer-grade devices reduce their capabilities for automatically and reliably detecting medical conditions. Given these limitations, cognitive personal informatics systems should be clearly framed as tools for exploration and enquiry, rather than as tools that automatically perform complex medical diagnoses.

Our findings suggest that EEG is most likely to make sense when accompanying less ambiguous data from other tracking technologies. This is because it may gain a context for interpretation and encourage an exploratory and enquiring approach, rather than fostering a belief that ‘the answer’ (e.g. to depression, dementia, etc.) can be found at the press of a button. Previous work (e.g., Ayobi et al., 2017), has similarly highlighted the importance of avoiding deterministic diagnoses and advocated exploration in self-tracking technologies for managing health and well-being.

Users were keen for some of the burden of data exploration and analysis to be reduced, for example by allowing the machine to select aspects of the data that a user may wish to pay particular attention to. Future work that aims to develop methods for directing users to particularly interesting ‘cuts’ of the data (Epstein et al., 2014), facets within data that are statistically correlated (as in S. L. Jones and Kelly (2018)), or important’ subsets of data, e.g. highlighting data that is ‘out of the norm’, may encourage users to engage with data, whilst simultaneously removing some of the burden of information overload that comes with high-dimensional data within a PI context.

#### **3.4.4 Challenge 4: Overcoming misconceptions & lack of understanding about the brain**

It was noted that some of the concepts and measures considered by participants lacked solid scientific grounding, or were based on misinformation or pseudoscience. For example, one participant mentioned wanting to evaluate her “biorhythms”; a specific theory that people’s daily lives are affected by a rhythmic cycle, which has been shown to be invalid (Hines, 1998). Our study illustrates that such beliefs may shape users’ interpretations and evaluations of brain-related tracking data. Previous studies (e.g., Herculano-Houzel, 2002) have shown that the public poorly understands many aspects of information about the brain. For example, there are commonly held misbeliefs that humans only use 10% of their brain, and that brain activity depends entirely on the external environment (Herculano-Houzel, 2002). Furthermore, participants in our study used diverse language to describe similar concepts. For example, several participants referred to the notion of mental ‘strain’, while others used terms such as mental ‘stress’, or ‘load’. Future cognitive personal informatics systems should aim to cater for users that have only basic knowledge of the brain, ensuring that information is presented in a simple and understandable way, that the system educates and informs users where possible, and that concepts are presented using commonly understood nomenclature, to address the possibility of diverse interpretations by users.

### **3.5 Conclusion**

This exploratory study sought to better understand the potential use cases and motivations for a cognitive tracking technology. The initial interview results suggest that there is an interest in this technology, and users are keen to track and gain insights about their cognitive activity. Users suggested using cognitive personal informatics systems for monitoring medical conditions, gaining awareness of their general well-being, improving understanding of themselves and others, as well as supplementing existing tracking technologies. However, for this to become viable several challenges must be addressed, namely: dealing with the indirection between meaningful psychological states and brain activity data, designing features to support diverse tracking styles, encouraging exploratory and enquiring approaches for dealing with brain data, and overcoming misconceptions and lack of understanding about the brain. Future work should aim to address two notable limitations of this study. First, our findings are based on initial reactions to a basic prototype, designed to prompt considerations about future CPI systems. We therefore lack longitudinal data and feedback based on use of a CPI system over an extended period. Second, there was limited functionality

within our prototype app. A system that is more fully developed, for example enabling more advanced interactions with the data, would likely provide additional and deeper insights into specific design considerations for CPI systems.

In summary, we find that there are several challenges to be addressed first before CPI can become a reality:

- Challenge 1: Addressing the indirection between meaningful psychological states and brain activity data
- Challenge 2: Supporting diverse tracking styles
- Challenge 3: Encouraging exploratory & enquiring approaches
- Challenge 4: Overcoming misconceptions & lack of understanding about the brain

We further address and explore these challenges as described in the following section, throughout the remainder of this thesis.

### 3.6 Addressing These Challenges

As described in Section 3.4.1, there is a need to address how people make sense of the presented data that comes from the EEG sensors and making a connection to meaningful state that they can interpret. Our current evidence suggests therefore that raw EEG data is likely to be of little benefit to the ordinary person in seeking to understand more about themselves when it comes their brainwaves. Similarly, although certain EEG wavebands may have correlates with specific brain activities, outside of research these are likely to be difficult to interpret. We therefore need to understand whether providing users with processed values, such that the raw values have already been categorised into an associated states (e.g Mental Effort), aides them in more easily understanding, interpreting and interrogating the collected EEG data.

In our further studies we aim to gain a greater understanding of what benefit there may be to providing users with these process values, referred to as metrics, in the studies described in the following chapters by utilising the values that are made available via NeuroSky's API library and limit the inclusion of raw EEG values and wavebands. An exception to this is in Study 3 where Alpha values are reintroduced due to its correlates with fatigue, which is a symptom of the conditions the population recruited to take part in the study live with. We will therefore be seeking to understand if there is a benefit to the inclusion of the metrics, or if there are further issues encountered with them.

In our next studies we aim to gain a better understanding of the types of goals that people might use when tracking data from novel sensing technologies. How do they determine what goals to set and how do they determine if they have reached these goals?

How does this differ between different styles of tracking? Do documentary, diagnostic and directive styles of tracking have different approaches to understanding the data from novel sensing technologies? Can anything be learned from these style be applied to the benefit of other styles of tracking?

To address the challenge of supporting diverse tracking styles applications may need to incorporate a variety of ways to record and review a user's data. For example, people who may be categorised as documentary trackers likely need less ability to interrogate and investigate their data as those who may wish to track for diagnostic reasons. However, documentary trackers may still be interested in simplistic ways of reviewing their tracked data.

We will also seek to better understand what additional features may be of benefit to encourage users to thoroughly explore and investigate their data. In particular it may be beneficial to highlight interesting pieces of information to people rather than expecting that are initially aware and able to find this insights themselves. This is perhaps important due to the limited or varying levels of knowledge regarding the data being collected from a novel sensing technology.

As the research conducted as part of this thesis is not seeking to specifically understand the brain and is looking to providing answers relating to novel sensing technologies more generally, we do not seek to address specific methodologies for overcoming misconceptions about the brain, as they outside the scope of the research. However, misconceptions about the data collected from novel sensing technologies is investigated further in Chapter 7. In this study we are seeking to determine whether data visualisations that more closely matches a person's preconceived expectation of how the data should look influences the amount of trust they may have for a specific device's reliability and accuracy. Such factors could possibly influence, change or reinforce people's misconceptions about the data they are recording from novel sensing technologies.

The ways we have sought to approach and answer these questions are discussed in the following chapters.



## Chapter 4

# How People Use a Novel Sensor PI System

## 4.1 Chapter Overview

In this chapter we discuss the results of a three-week in-the-wild study of a multifaceted PI system, which was conducted to answer some of the questions raised by the previous study. In this study we attempt to better understand how people make sense of data produced by ‘novel sensor technologies’, characterised as those which are wearable, produce large amounts of data with which people may not be familiar, trust in the data may vary or be difficult to determine and transparency of how the data is produced may be lacking. The aim of this study was to answer the following specific research question:

**RQ1** How do users record, collect, analyse and make sense of data from ‘novel sensor technologies’ using a multifaceted PI system?

**RQ1.1** What difficulties do users have in making sense of sensor data?

**RQ1.2** How do users appropriate the data from the tracking application - what visualisations, records, additional analyses, etc. do they create and why?

**RQ2** What are the implications of providing people with metrics which have not yet been clinically validated?

**RQ2.1** Does the use of abstracted metric values rather than EEG wavebands have an effect on self-understanding/self-reflection/self-improvement?

In addressing **RQ1** we seek to better understand the way participants approach the recording and analysis of data using a multifaceted PI system, in combination with the BCI headset. **RQ1.1** specifically looks at the difficulties people encountered when attempting to make sense of the data recorded. **RQ1.2** aims to drive our understanding of the additional ways in which people attempt to analyse the data, beyond the use of the functionality provided in the implemented application used by participants, to provide insights for design considerations of future PI systems.

For **RQ2**, we are looking to understand the implications of providing metrics which do not yet have clinical validation but have initial experimental suggestions of presence. Some of the metrics provided by the NeuroSky headset are based on published academic papers which suggest that certain values or responses can be measured via EEG. However, the NeuroSky headset itself has not been clinically validated as a tool capable of measuring these responses. In contrast to the way values were presented in our initial study, **RQ2.1** specifically looks to understand the benefits and disadvantages of providing metrics rather than raw data values to participants.

## 4.2 Methodology

To address these research questions we conducted an in-the-wild study over a 29-week period with 10 participants, each participant spent three weeks taking part in the study. Data was collected by way of qualitative interviews. Thematic analysis (Braun & Clarke, 2006) of post-study interviews responses was used to understand the ways in which participants used the applications, difficulties they encountered and the ways in which they wanted to use the applications but were unable to in the current implementation. Ethical considerations for this study followed the Department of Computer Science’s 13-point checklist, see Appendix G. Data pertaining to this study is retained in accordance with the data management plan found in Appendix F.

### 4.2.1 Mobile Application

As part of the in-the-wild study participants were provided with a mobile application capable of recording and presenting data from a NeuroSky headset, Fitbit for heart rate, Spire for breathing rate, as well as manually logging mood data and perceived workload information. The participants were therefore able to undertake the recording and investigation of personal informatics data from a variety of sensors, in addition to manually captured data, in one single application.

The initial features included in the application were based on the results of the initial exploratory study of 16 participants that gathered their initial thoughts about tracking data from novel sensors, as presented in Chapter 3. In particular they were asked to consider how they might use and interrogate data recorded by an EEG headset, namely the NeuroSky MindWave. This included the ability to track data over multiple sessions and to compare data between recording sessions. Features were implemented taking direction from participants’ responses and inspiration from the design of existing PI-style apps, such as Fitbit.

We developed a mobile application which allows users to collect and record EEG data, as quantified metrics, from a NeuroSky headset. The metrics recorded and presented by the application were extracted from the headset’s API. Additionally, the application synchronises heart rate data from a Fitbit<sup>1</sup> smartwatch and breathing rate data from a Spire Stone<sup>2</sup>, for each EEG recording session. The application also incorporates manual mood logging via an Affect Grid (Russell, Weiss, & Mendelsohn, 1989) with a free-text area, and NASA TLX (Hart & Staveland, 1988) ratings for perceived workload.

---

<sup>1</sup><https://fitbit.com>

<sup>2</sup><https://spirehealth.com>

Metric	NeuroSky Description	Values
Alertness	The moment-to-moment level of alertness or vigilance of a user. High Alertness values indicate the user is in a state of focus, while low values represent a relaxing state of mind.	Range: -1 to 1 BCQ Valid: 0 or 1
Appreciation	The moment-to-moment level of enjoyment or appreciation of a user towards an external stimulus, e.g. video, music, etc, based on the user's brainwaves.	1: Not at all 2: Low 3: Medium 4: High
Attention	How focused or single-minded the user is at the moment.	1-20: strongly lowered 20-40: reduced 40-60: neutral (baseline) 60-80: slightly raised 80-100: elevated
Cognitive Preparedness	The moment-to-moment capacity of a user's brain for optimal cognitive performance on a relatively complex task. In other words, it represents the brain's capacity for higher level cognitive functions.	Range: -1 to 1 BCQ Valid: 0 or 1
Creativity	The moment-to-moment level of activity underlying creative cognition of a user. High Creativity values indicate stronger brainwave activities promoting innovative and creative thinking.	Range: -1 to 1 BCQ Valid: 0 or 1
eTensity	The moment-to-moment intensity of a user's emotions towards an external stimulus.	1: Not at all 2: Low 3: Medium 4: High
YinYang	The moment-to-moment emotional response of a user towards an external stimulus.	-1: Unpleasant emotion 0: Neutral emotion 1: Pleasant emotion
Familiarity	The moment-to-moment learning progress experienced by a user while practicing a new skill. It gives an indication of the "Learning Curve" during the skill acquisition process.	Progress: 1 to 5
Meditation	How calm and clear-minded the user is at the moment.	1-20: strongly lowered 20-40: reduced 40-60: neutral (baseline) 60-80: slightly raised 80-100: elevated
Mental Effort	The moment-to-moment mental workload experienced by a user during a task. The harder the user's brain is working at the time (ostensibly on the task), the higher the Mental Effort value.	Total: 720000 to -720000

Table 4.1: NeuroSky Metric Descriptions

The NeuroSky MindWave headset has an API that allows measuring of the metrics, see Table 4.1, for each session these values were summarised as described in Table 4.2.

Metric	Visualised Data	Visualisation Type
Alertness	Proportion of time spent in ‘reached’ threshold.	Pie Chart
Appreciation	Proportion of time spent in range 1-4.	Pie Chart
Attention	Proportion of time spent in ranges.	Pie Chart
Cognitive Preparedness	Proportion of time spent in ‘reached’ threshold.	Pie Chart
Creativity	Proportion of time spent in ‘reached’ threshold.	Pie Chart
eTensity w/ YinYang	XY plot likened to circumplex model.	XY Plot
Familiarity	Rate of Familiarity over time	Line Chart
Meditation	Proportion of time spent in ranges.	Pie Chart
Mental Effort	Total ME over time	Line Chart

Table 4.2: NeuroSky Metric Visualisation Types

#### 4.2.1.1 Session Recording

Participants recorded their data in sessions in a way very similar to the previous version of the application used in the exploratory study. Participants were able to add tags to recording sessions so they would have some contextual labels as a point of reference when they were reviewing their data. Participants were able to see of a live view of the data being recorded on a line graph as the session progressed (see Figure 4-1a). Participants were then able to pause or stop recording once they felt they had recorded sufficient data. After participants had finished recording their session they were able to record a mood rating using the Affect Grid (see Figure 4-1b) and were then asked to complete a NASA TLX assessment (see Figures 4-1c & 4-1d).

#### 4.2.1.2 Session Review

Figure 4-2 shows how these EEG metric summaries were displayed for users on a per session basis. Participants could access averaged data over multiple sessions for a metric by selecting that metric’s graph. Figure 4-3 shows how Fitbit heart rate and Spire breathing rate data were displayed, as a line graph with the time spanning from 30 minutes before a recording session’s start time to 30 minutes after a session’s end time. Also seen in Figure 4-3 are the Affect Grid rating summaries and the NASA TLX summary that was accessible from the session overviews.

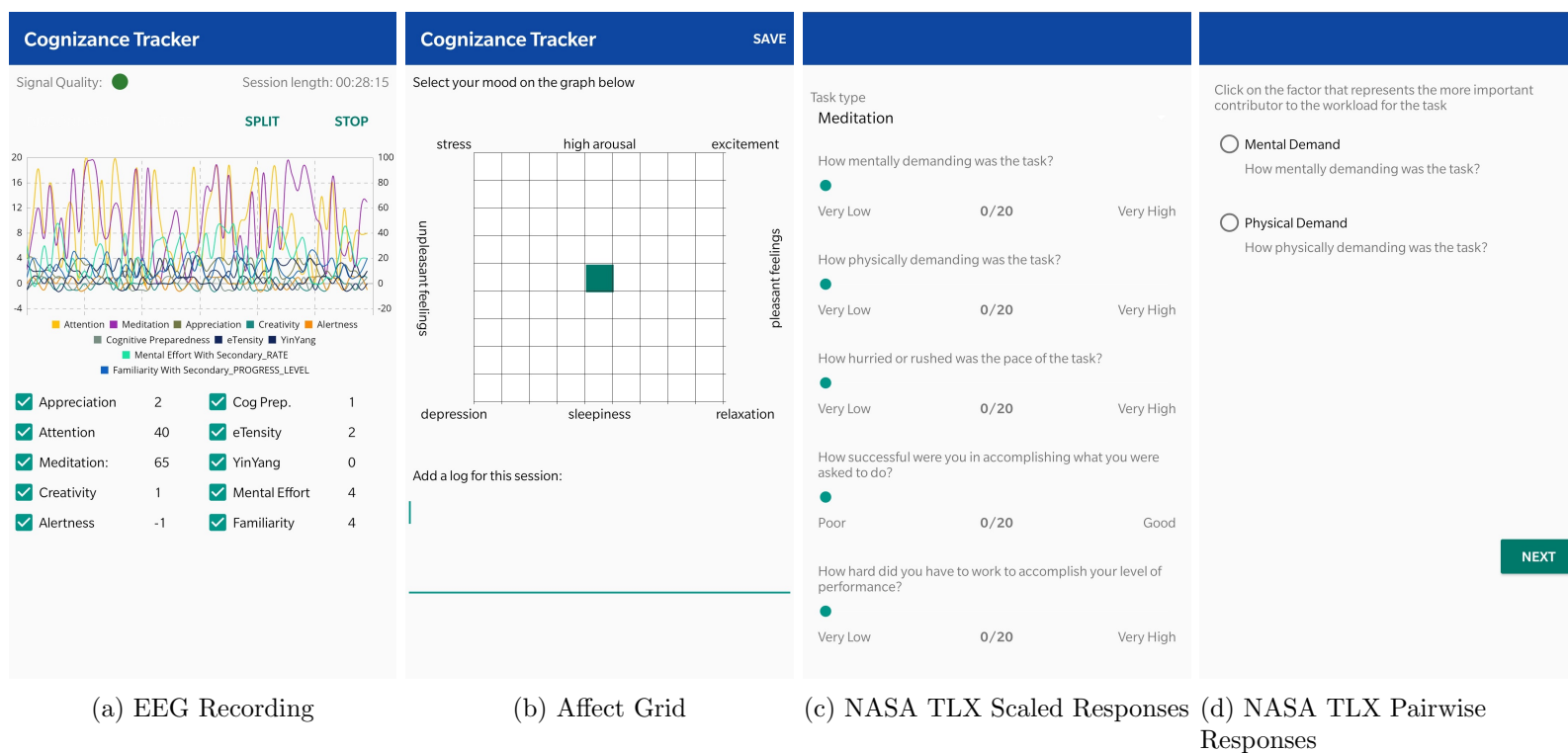


Figure 4-1: Session Recording

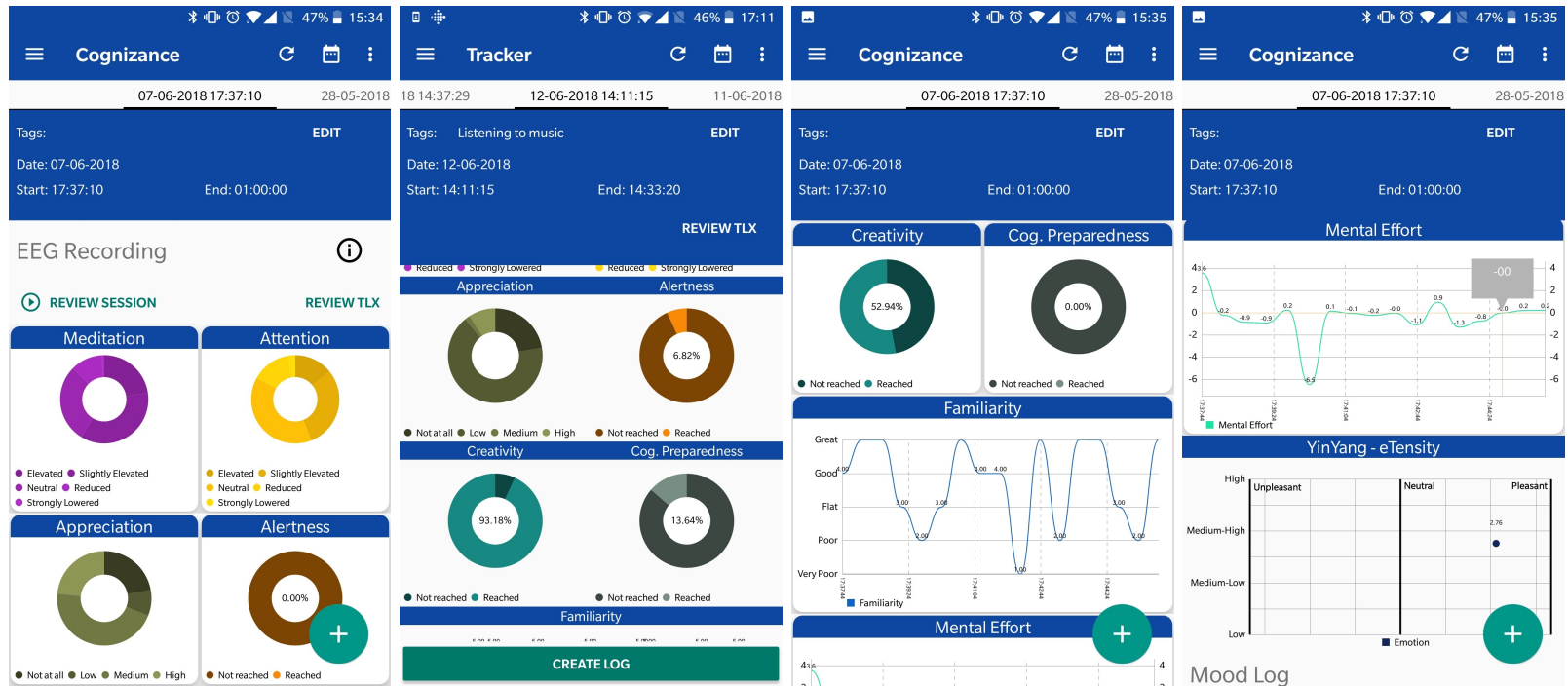


Figure 4-2: Session Overview - EEG Metric Summaries

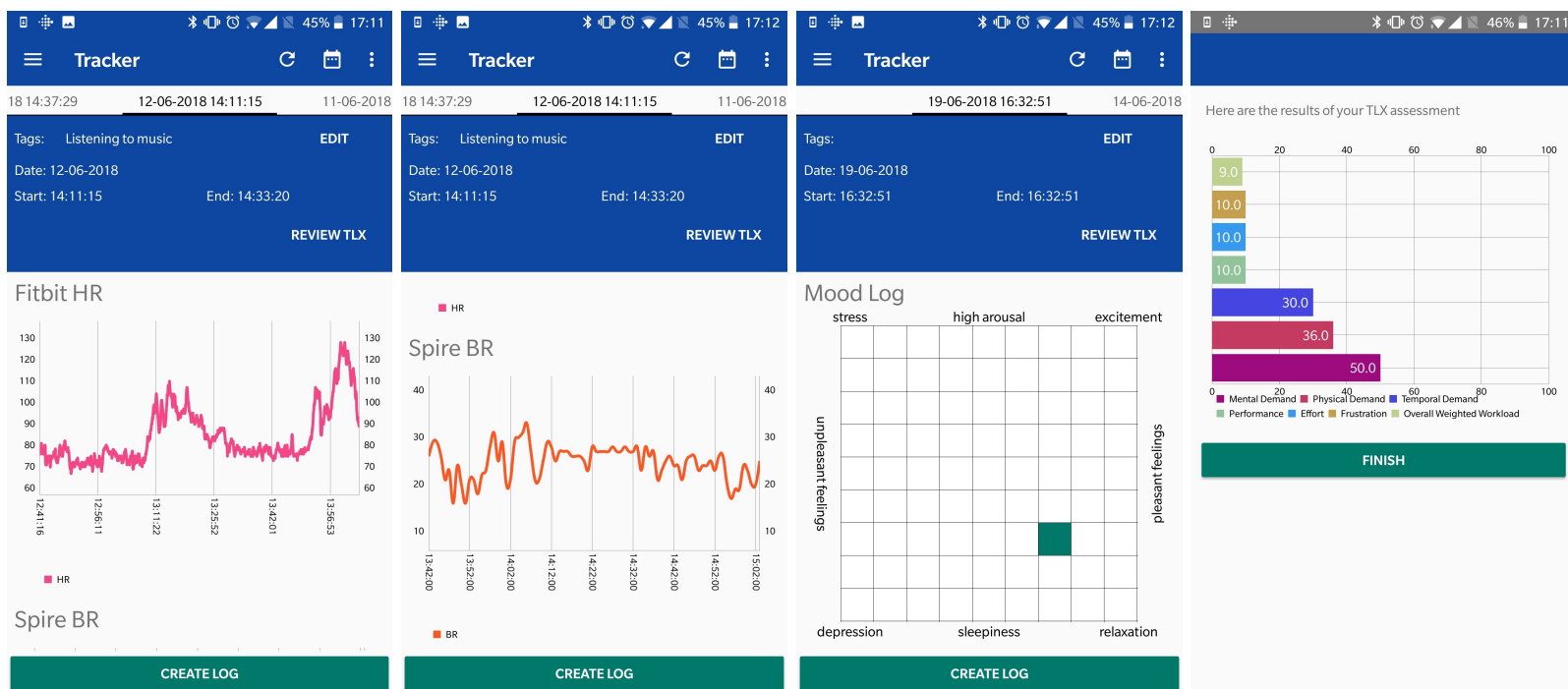


Figure 4-3: Session Overview - HR, BR, Mood and NASA TLX Summary



### 4.2.2 Participants

Participants were recruited via email, online- and non-virtual noticeboards at the University of Bath. Participants were required to fill in a pre-screening questionnaire to gather some demographic as well as data about previous tracking experience. Pre-screening gathered data from 31 participants (13 M, 17 F, 1 PNA), with ages ranging from 19 to 63 (M: 31.77, SD: 11.5). Two participants were not contacted as they were located in a different country, which made loaning the hardware impractical. Three participants were excluded for ethical reasons; we opted to exclude those with previous diagnoses of a mental health condition as we were unsure how the application and metrics may be interpreted and did not wish to unintentionally negatively impact them. The remaining participants were then emailed a copy of the information and consent form (see Appendix C) and asked to confirm their willingness to participate having read this. Of the remaining 26 participants, 10 participants (6 M, 3 F, 1 PNA) agreed to attend a briefing session, with ages ranging from 19 to 58 (M: 28.30, SD: 11.72).

Participant #	Age	Gender
P1	27	Male
P2	36	Male
P3	32	Female
P4	58	Male
P5	19	Male
P6	22	Male
P7	21	Prefer not to say
P8	21	Male
P9	23	Female
P10	24	Female

Table 4.3: Study 2 - Participant Demographics

Four participants had previous experience with PI technologies for self-understanding or self-improvement, ranging in periods from less than a month to >1 year but less than 2. Three had used wearable tracking devices previously, two had used mobile applications, and one had used a Muse headband. Participants' previous reasons for self-tracking included weight control, sleep, step tracking, as well as to "increase focus" by the participant who had previously used the Muse headband. Four participants had previous experience with using mindfulness technologies, with one having less than six months experience, and the remaining had less than one months experience. Five participants had experience of non-technological mindfulness practices ranging from 5+ years (2), <3 months (2), <1 month (1), including practices such as yoga (1), meditation (3), and tai-chi (1).

Two of the recruited participants did not use an Android phone and were provided with one to use for the duration of the study with the Cognizance Tracker application installed. All of the other applications were installed on their main phone, such as the Spire and Fitbit companion apps.

Participants were compensated with a £5 (\$6.50) Amazon voucher for their participation in the post-study interview and were entered into a draw to win one £50 (\$65) Amazon for their overall participation.

### **4.2.3 Procedure**

Participants were asked to make a daily recording session: this involved wearing the MindWave headset, Fitbit watch and Spire Stone, whilst performing a task of their choosing for roughly 30 minutes. The participants were asked to make a minimum of 7 recording sessions over the 21 day period to be eligible for the post-study interview. This was to ensure that participants had sufficiently interacted with the technologies to be able to discuss their usage during the post-study interview. No participants were excluded at this stage.

During the study briefing sessions, lasting about 30 minutes, the Cognizance Tracker application was explained to the participants. They were provided with assistance installing and setting up Fitbit, Spire, Cognizance Tracker and each application's associated hardware: Fitbit Versa, Spire Stone and NeuroSky MindWave respectively. Time was spent with each participant to explain how to use all of the hardware. Before the briefing ended participants had made a short recording session of a few minutes, to ensure they understood how to place the sensors and that they understood the what was displayed during a recording session. Participants were also encouraged to reach out to the researcher if they had any issues during the study period, technical or otherwise.

### **4.2.4 Interviews**

During the post-study interview participants were asked for their initial comments about the technologies used over the three-week period with the hardware. They were then asked about their thoughts on recording EEG data, what sort of questions they felt their EEG data could help them to answer and if they felt they gained any benefit from recording their EEG data. Participants were asked whether they found any relationships between the data they were collecting and whether any of the data stood out or intrigued them. Participants were asked what kinds of tasks they carried out during

their recording sessions, what tasks they might consider in the future and what tasks they wanted to perform, but were unable to.

Participants were also asked specific questions to understand how they used the application, such as how they used the application to determine if they reached specific goals or values. Participants were asked how many sessions they logged, and if this was less frequent than daily they were asked what would have encouraged them to use the device more and why they stopped using the devices.

Participants were asked about data accuracy: how accurate they believed the data was, whether the accuracy of the data affected how they might use the application and devices for tracking, and whether they felt the data may be more useful if additional data was collected, i.e. would recording additional session make the data more useful? Participants were asked which features of the application they made most use of and whether the way the data was presented was useful, whether there was too much or too little data shown, whether they trusted the data they were being shown, and what features would encourage them to make more use of the application.

Participants with previous PI experience were asked how tracking using the Cognizance Tracker application compared to their previous experiences tracking data.

Participants were then asked what other data sources they might like to combine with EEG data to learn more about themselves, how they might use a similar technology in the future, how long they might spend recording EEG data, the locations in which they are likely to use such a technology, who they believe may benefit from recording EEG data, their concerns about and demerits of recording EEG data, and whether a similar emerging sensing technology would have a positive or negative impact on their life.

The interview responses were analysed using thematic analysis (Braun & Clarke, 2006). The interviews were transcribed in full, a total of 04:11:00 of audio was transcribed (M=00:25:08 per participant, SD=0:07:24). After reviewing and re-reading the transcribed responses, codes were generated inductively; directed by the data. Once the data was coded, the codes were reviewed and grouped into themes of shared meaning. These themes were named and defined in relation to the relevant research questions.

### **4.3 Interview Results**

The following sections present the results of our thematic analyses in relation to this chapter's aforementioned research questions. A summary of the research questions,

related themes and sub-themes are presented in Table 4.4.

#### **4.3.1 How do users record, collect, analyse and make sense of data from ‘novel sensor technologies’ using a multifaceted PI system?**

The following themes of *Willingness to Record with Novel Sensing Technologies*, *Worries About Recording with Novel Sensing Technologies*, and *Scheduling and Timing* describe the difficulties, concerns and time-management issues that participants encountered in recording and collecting data from a novel sensing technology, respectively. The theme of *Variation of Tracking Activities* describes the vast spectrum of situations in which participants used, or would like to use, novel sensing technologies to learn more about themselves.

##### **4.3.1.1 Willingness to Record with Novel Sensing Technologies**

Although participants evidently have some willingness to use novel sensing technologies, having chosen to take part in the this study, they still expressed some issues in relation to the technologies they used. The device that participants found most problematic was the NeuroSky MindWave. The main factors encountered appear to centre around sub-themes of the *Perception Of Others/Use in Public*, being *Restricted by Sensing Technology*, and the *Novelty of Technology*.

##### **Perception of Others/Use in Public**

One of the factors that affected participants’ willingness to record data using the NeuroSky MindWave was their perceptions about what other people would think about their use of the device. This was expressed as an unwillingness to record using the headset in public, “*Most of the time it was at the end of my day because I was quite embarrassed to wear it in front of other people*” (P3), or in front of people who did not understand what they were doing, “*when I’m sat at home working and it’s only my wife around so it’s only her wondering what on earth I’ve got on my head*” (P1).

However, this was not an issue shared with all participants, with some stating that they had in fact used the headset in public, “*I do not mind wearing it throughout my day in the university or on my way travelling*”. Participants suggested that this type of technology would be more acceptable if it were to become less noticeable, “*I would’ve appreciated if it was quite so obvious, and I would have had to explain everything to every person I met*” (P7).

These comments highlight that a vast majority of users desire novel sensing technologies

Theme	Sub-theme
<b>4.3.1 How do users record, collect, analyse and make sense of data from ‘novel sensor technologies’ using a multifaceted PI system?</b>	
Willingness to Record with Novel Sensing Technologies	Perception of Others/Use in Public Restricted by Sensing Technology Novelty of Technology
Worries About Recording with Novel Sensing Technologies	
Scheduling and Timing	
Variation of Tracking Activities	
<b>4.3.2 What difficulties do users have in making sense of sensor data?</b>	
Insufficient Knowledge	
Amount of Data Shown	
Gaining Insights From Data	
Effect of Live View During Recording on Analysis	
Preferring Specific Types and Sources of Data	
<b>4.3.3 How do users appropriate the data from the tracking application - what visualisations, records, additional analyses etc. do they create and why?</b>	
Analysing Data	Comparisons Correlations Determining Factors Goal Setting Trends Summaries
<b>4.3.4 What are the implications of providing people with metrics which have not yet been clinically validated?</b>	
Trust	
Accuracy	
<b>4.3.5 Does the use of abstracted metric values rather than wavebands have an effect on self-understanding/self-reflection/self-improvement?</b>	
Difficulty Interpreting Metrics	
Data Meeting Expectations	

Table 4.4: Summary of Themes

that allow them to discreetly record their tracking data. Wearability alone is likely not a sufficient factor to foster using such a technology in everyday life, rather users want to be able to track information without also drawing attention to the fact that they are doing so. Pfeiffer, Von Entress-Fürsteneck, Urbach, and Buchwald (2016) suggests that ‘aesthetic’ does not play a significant role in adoption of self-tracking technologies, nevertheless, participants’ responses in this study suggest that it does play a role depending on how discreet that technology is. Although the appearance of the technology may not be important, the aesthetics of a device still requires discretion.

### **Restricted by Sensing Technology**

Another factor was that the headset prevented participants from collecting data about themselves in all of the contexts that they would have been willing to do so. For example, participants expressed wanting to have recorded data during a drive (in the car), or during physical activities, but found it difficult or impossible to do so. Participants highlighted that with the current design of the headset *“it’s not realistic to wear it when you’re doing anything energetic”* (P4) and that they *“would love to have known how the data was different, say when I was out for a run, but that just didn’t seem practical to gain that”* (P2). P8 similarly expressed desire to record their EEG data in more active scenarios, *“I would love to wear it, in a more locomotive kind of setup, running or gym, playing sport because I would like to test the physical stuff”* (P8).

These comments highlight that designers of novel sensing technologies must consider the ways in which users are likely to want to record data, and thoroughly consider whether the design is appropriate in particular during activities other than just sedentary tasks.

### **Novelty of Technology**

One of the things that encouraged engagement with tracking appears to be the novelty of the devices, *“Obviously because it’s quite a novelty with the headset, because it’s not something I’ve used before, I was kind of experimenting more so than going deep into trends”*. References to the devices’ novelty was seen across the range of devices that participants had used. P1 had never used a smartwatch before, *“I never really tried a smartwatch, or anything like that, so I found that most helpful.”* (P1). Whilst another participant found that the Spire was novel, *“It was interesting in the first case with the Spire thing to see that such a thing even existed. I hadn’t conceived of being able to detect breath patterns that way.”* (P4).

A related factor, therefore, is whether once this novelty wears off, users are less likely

to engage with tracking. One participant stated that they used the headset less as time went on, *“Initially I was a lot more prone to using it because I was interested in technology but as the days went by I felt that I would rather just use the watch at the end of it.”* (P9).

This suggests that the novelty effect of these devices reduced the likelihood of continued use. Previous research into the effects of novelty of activity tracker usage highlights that this effect often lasts for about three months and that other extrinsic and intrinsic factors are likely to be the reason for continued use beyond this (Shin, Feng, Jarrahi, & Gafinowitz, 2019).

#### 4.3.1.2 Novel Sensing Technology Concerns

Worries about recording with novel sensing technologies centred around a few main issues. Firstly, participants were concerned about data protection and who might be viewing their data. P3’s concern was that they did not know where the data was going, *“At the beginning I felt like a robot and I didn’t know where this data is going. So that was a bit worrying.”* (P3). Whilst P4 was more concerned that wherever the data was going, it should be *“appropriately protected”* (P4). P9 expressed their concern about how a company might appropriate their data, *“I would be considering questions such as why does this company need to know how I feel”* (P9).

The second concern participants expressed was that their data might be used for nefarious reasons, *“Conspiracy theory, Big Brother trying to control people’s minds maybe. Maybe it’s not going to happen but it’s the next step.”* (P3). This concern appears to be exacerbated by recent concern about how social media companies use data,

*It’s sort of like the Facebook thing where they were analysing your posts to figure out your emotions and advertising based on that it would be similar to that because if they could figure out what your feeling, what your current mood is, what your overall mood is, then I wouldn’t want them to know that really. And if they could figure out what things I appreciate and what makes me feel meditative, I wouldn’t want them knowing that ever. I think it would mostly just be information I keep to myself or share with people willingly, but never analysed beyond that.* (P7)

Thirdly participants were concerned that people may become ‘slaves to the data’ or that these technologies may be a solution to a non-existent problem,

*It all depends on how you use it. So you shouldn’t always try to find a*

*use for something because it's already there. So for me the first thing is to indicate what your goal is and then see if that technology can help you with that and not just, "Oh, let's do that." because it's something new, and it's a new toy. So the other way round. (P3)*

Finally, one participant suggested a concern related to being unsure about what impact recording EEG data might have on them physically, *"I'm not sure how the EEG works, does it disturb my brainwaves or not. I'm not sure...when thinking about the physical thing, we have the battery or the current that runs through [it]" (P6).*

Pfeiffer et al. (2016) also suggests that trust in the vendor of the device (e.g. NeuroSky) also plays a significant factor in self-tracking technology adoption. Responses here also suggest that users are less likely to use technologies if they believe that the vendor's intentions with their data are not altruistic.

#### **4.3.1.3 Scheduling and Timing**

Participants described scheduling recording sessions in a variety of ways. Some participants stated that they found they needed to purposefully set aside time to record data, *"I literally, artificially, or actually dedicate the session times so I actually record them, so I didn't actually wear it and go for grocery shopping to Sainsbury's" (P8).* Whilst others suggested if they were to continue using the technologies they would schedule recordings, *"I might try to... Initially try to schedule a regular slot and adhere to that more rigorously." (P4)* as during the study they just recorded when they remembered.

Some participants felt that recording sessions interrupted their normal activities. P1 felt that the Spire's feedback was annoying due to it advising them that they were tense, *"It was getting a bit annoying when like playing computer games and it was telling me I was tense" (P1).* Whereas P8 felt that recording was slightly impractical and required adjustments to be made to enable them to record data,

*I couldn't take off my jumper wearing that, so I literally have to make sure I'm not wearing anything or that I do not have to take off anything while wearing that. So I remember once I was wearing a jumper when I was recording and I was drinking water and a bit of water fell so I wanted to take off my jumper but it was in the middle of the recording so I literally couldn't. (P8)*

Regarding reviewing data participants mentioned that they felt that it would be be-



neficial to have “*some sort of routine (for) reflecting on the data and using it as some sort of quantitative metric.*” (P1), such as having “*a review session once a week or a month and your goal*” (P1).

Prior work suggests that when self-monitoring, recording every instance of a variable under investigations increases alterations in performance or behaviours (Korotitsch & Nelson-Gray, 1999), thus creating schedules around recording and analysing data is likely to be beneficial by ensuring that users are tracking relevant data. PI tools therefore should encourage and find ways of engaging users in tracking their data and scheduling data collection.

#### 4.3.1.4 Variation of Tracking Contexts

Participants mentioned having or wanting to record a variety of different contexts, highlighting the variety of situations in which people want to collect data about themselves. Presented in Table 4.5 are the occurrences of activities mentioned by participants during their recording sessions.

Context	No. Participants
Cooking	5
Reading/Editing on Screen	4
Studying	3
Watching Movies or TV	3
Working/Admin	3
Chores	2
Eating	1
Exercise	1
Meditating	1
Playing Games	1
Various Locations	1

Table 4.5: Variety of Recording Contexts

Participants’ responses suggest that the contexts in which they are likely to want to record are numerous and diverse, varying between ‘sedentary’ tasks, low and high impact ‘physical’ activities, and captured in various locations, with participants keen to determine differences in data recorded in different contexts and locations.

#### 4.3.2 What difficulties do users have in making sense of sensor data?

Difficulties that participants had when making sense of the sensor data presented in the Cognizance Tracker application are reported here under the themes of *Insufficient*

*Knowledge or Information, Amount of Data Shown, Gaining Insight from Data, Effect of Live View During Recording on Analysis, and Preferring Specific Types and Sources of Data.*

#### **4.3.2.1 Insufficient Knowledge or Information**

One of the main issues that participants experienced in attempting to make sense of their data was that they felt they lacked sufficient knowledge to interpret that data. Participants stated that they did not understand the metrics, *“I feel like I didn’t fully understand the metrics it was giving me”* (P10), or that the metrics were too technical, *“Probably in a bit more non-technical and broken down jargon, because not everyone would understand what “YinYang” or those technical terms are”* (P8).

One participant suggested that they required neuroscience knowledge to be able to make sense of the EEG data, *“I think I need to know more about neuroscience to actually understand it. But I kind of got a vague... No. No I don’t deeply understand what it was showing me”* (P5). Whilst another suggested that perhaps a third-party could analyse the data on their behalf, *“I’m sure there are applications where you can take the data and present it to someone who perhaps knows a bit more than you do. ... I’m not necessarily thinking like a doctor, maybe somebody who assesses your learning or your focus”*.

Other participants suggested that less onus be placed on the user to have to interpret the data, such as comments made by P2, *“I’d almost want something quite simplistic more like the Fitbit app where it can sort of interpret the data and tell you in a few simple words or graphs rather than sort of having all of that there to have a look at if that makes sense?”* (P2) and P3,

*I don’t have to look at everything at once and try to connect everything at once. If there was somehow it tells me, “Oh you were focused.” or “You reached the focus level based on a few metrics” and then if I want I would go deep and look at the specific values.* (P3)

This suggests, in relation to novel sensing technologies, that there is a need to provide users with sufficient knowledge to begin their interpretation of the data. This is perhaps particularly important when the data being recorded is not something that one might consider to be more ‘general knowledge’ such as the implications of heart rate data. Ding, Wei, Gui, Gu, and Zhang (2021) has similarly highlighted that a lack of understanding about what values are being measured (e.g. psychological vs. physiological stress) and therefore how to interpret the data can lead to mismatches in understand-

ing presented data.

#### 4.3.2.2 Amount of Data Shown

Participants generally considered the amount of data shown to be either sufficient or too much. No participants commented that there was insufficient amounts of data shown. For the participants that felt that there was too much information shown, they became overwhelmed or confused by the amount of data, *“I knew about the Spire device, and Fitbit and the headset in general but then in the app when there were all these metrics I got confused somehow with them.”* (P3). Others felt that there was too much information to start with, *“it’s providing probably rather more information but that’s probably a good place to be initially. More information than I could digest or knew how to digest”* (P4), suggesting that extra information might be useful but they would only use it more if they could more easily find the information that was of interest to them, *“Using it more would mean being able to get the message from the data more readily than one can at the moment”* (P4).

For those who felt there was sufficient data they highlighted that they were able to use the colour coding of data to make effective comparisons, *“it’s not too much, and we can compare from the colour, it’s very nice to see of different contrasts of each topic”* (P6). P5 felt that they could always show and hide different pieces of information depending on what they wanted to see, *“It’s not like you were getting overloaded with... You could always just take on and off the different things so if you just wanted to look at.”* (P5).

These two competing thoughts about the amount of data shown highlight an interesting point. Some participants felt that it was sufficient to be able to hide and show pieces of data on a case-by-case basis, when it was of interest to them. For some users, who felt that there was too much information, it may be necessary to provide more prominent cues on how to remove data that is not of interest.

#### 4.3.2.3 Gaining Insight from Data

Some participants’ comments suggest that recording EEG data increased their self-awareness rather than providing specific insights, *“Just by wearing it you’re conscious of wearing it, so probably would influence how you’re thinking when you wear it.”* (P7). Although increasing self-awareness may be beneficial to some people, the comment also suggests that users perceive a potential negative impact of cognitive activity (and associated data) being altered by virtue of wearing the devices, which may make it more difficult to analyse and extract insights from the recorded data.

P10 said that they would not be confident making decisions based on the captured data, suggesting that they needed reassurance, from someone with more knowledge, that the sensing technology was working as expected, *“probably I just need a reassurance that whatever is happening with the Spire is fine. Probably from a more experienced person or the developer or whoever”* (P10). This comment suggests that users require evidence that data is being captured accurately before they are willing to make decisions based on the data that is being presented.

Other participants suggested that they had just used the application to ‘get’ data and had not really attempted to analyse or make sense of the captured information, such as comments made by P1 *“I was mainly just logging sessions and briefly looking at it.”* (P1) and P5, *“But at the moment it was just getting data.”* (P5). This type of tracking is reminiscent of the documentary style of tracking highlighted by Rooksby et al. (2014), whereby people track data to document the things they have done, rather than to change or alter things which would require analysis of the data.

P9 suggested that they would be more inclined to use this type of technology for immediate intervention rather than reflection,

*I feel the stone helped with my mood only because I would end up getting anxiety whether it was in class, or whether it was about a coursework and it would just be like, “OK, I need to tone down.” So I think those two [watch and stone] especially really did impact me a lot more.* (P9)

This highlights that some people, although they are willing to record and have data about themselves analysed, do not wish to perform such analyses themselves. Similar to documentary style trackers, they are willing to record the data, but participants responses suggest they would rather have a device alert them to new insights or information of interest and provide suggested actions to take on the basis of the data.

#### **4.3.2.4 Effect of Live View During Recording on Analysis**

Although some participants said that they valued the live view of their data during recording, *“I would say the chart itself, the live reading that was very interesting”* (P9). Another participant highlighted that although the live view was interesting, it could be distracting. The participants felt that this could result in the recorded data not being a true reflection of what they were really doing during a recording session because it would include information as a result of having reviewed the live view in the moment,

*[I]t was interesting seeing the live feed as well because you can see when*

*you're creative, you can see it live. But I guess you get distracted and then you look at it and then it's not quite the task at hand.* (P5)

Although these comments specifically refer to watching the live view of EEG data, this is akin to being able to check heart rate in the moment on a Fitbit device. In the case of heart rate, looking at a Fitbit during exercise does not necessarily detract from the effect of the exertion on the heart itself, assuming the person continues to exercise at the same time. However, one might consider that watching your heart rate rather than focusing on the physical activity itself might be a distraction that could result in less effort being expended, thus reducing heart rate by attending to the fitness tracker rather than exercise. Therefore the physiological signal being measured, and how quickly it reacts to an external stimulus, might be a factor that designers should account for when deciding whether a live view of data is an appropriate feature to include in a PI system.

#### **4.3.2.5 Preferring Specific Types and Sources of Data**

Participants in this study made use of three wearable sensing technologies and a variety of manual tracking means; participants' responses suggest that they tended to prefer specific sources of data. However, this appears to quite individualistic, with no specific source being mentioned more than others.

Participants suggested that part of the reason for this was that some data sources were more readily understandable (heart rate and breathing rate), *"Something like heart rate is quite... To me that seems more easily quantifiable rather than being meditative"* (P5), or obvious, *"the data it's [the Fitbit app] giving you is really obvious data as well"* (P2).

Familiarity with a specific type of data also appears to be factor that influences preference towards one source of data over another, *"I guess I've seen that data [heart rate] before so... I knew how [to interpret it]"*. One participant deferred to the Spire because it was 'more focused' with what it did, *"it was more focused, didn't have as much but what it did do, it did well."* (P7). This suggests both that users may tend to favour sensors which appear to be effective at capturing one measurement particularly well, rather than a sensor technology that does many things, but none of them particularly well.

Another preference for specific sources of data was due to the capabilities of different sensors, either via the device itself or the device's companion app. For example, P9 suggested that they preferred Fitbit as a source of data due to the apps ability to motivate them to take more steps. Whilst they preferred the Spire due to the in-the-

moment nudges it gives by vibrating.

*[T]hey're a lot more on the point and very momentary. Maybe if it's something of that sort then I would rely on how the stone would vibrate but if it was something regarding motivation or exercise I would rely on the Fitbit for that because it does a couple of boosts like, "Oh you've completed 10,000 steps." (P9)*

Further highlighting the previously mentioned preference for data sources that are able to give in-the-moment feedback rather than something that can be analysed later to change behaviours.

### **4.3.3 How do users appropriate the data from the tracking application - what visualisations, records, additional analyses etc. do they create and why?**

The theme of ‘*Analysing Data*’ encompasses the types of analyses that participants either suggested they had carried out, had seen during use of the application, or would like to carry out in future.

#### **4.3.3.1 Analysing Data**

The types of analyses that participants performed or wanted to perform based on interview responses appear to fall into the following six types: *Comparisons*, *Correlations*, *Determining Factors*, *Goal Setting*, *Summaries*, and *Trends*.

#### **Comparisons**

Comparisons were suggested as a way of comparing different ‘activities’. For example, participants wanted to compare the effect of physical activities on the metrics being recorded, *“I tried to experiment a bit with going for a run beforehand and things like that, to see if that would clear my mind”* (P1), comparing the effect of different movies, *“Like when I watch this video how does it affect my brain”* (P6), between time and situation, *“somehow try to compare it with the time and situation”* (P1), between tasks, *“When I was I was doing some work, there was a few times where I could see the actual difference from one task to the other”* (P2), and locations or settings, *“I mean if you’ve got the location data . . . I guess that would be really interesting”* (P1).

Participants suggested that these comparisons could be done between sessions as a whole, *“I see from the score when the recording is finished and I compare from the*

*previous recording how far it increases” (P6), as well as over time, “some graph showing the different scores across a week” (P1).*

Participants also suggested that comparisons could be done across people,

*[I]t would be very nice to record the various kinds of people in society . . . to do various kinds of social research like many kinds of things and then after when enough data, from enough samples are modelled that kind of particular kind of a thing. (P8)*

## **Correlations**

When asked if they had found correlations within the data they reviewed during the study, participants suggested they found correlations between a variety of different metrics such as sleep (Fitbit) and tension (breathing rate), *“I would have probably observed that I had less sleep and was more tense during the first week of term. Are we surprised?” (P4)*, concentration and mood, *“When I’m happy the concentration is high, but if it’s a bad day of coursework or anything else it affects the mood to concentrate on the work” (P6)*, Creativity and movie genre, *“if I watch a movie, an excited movie, my mental demand is quite varied but if I watch something that is funny the Creativity is very high and as my relaxation [sic]” (P6)*, Alertness and Attention, *“I noticed if I’m more tired then my alert levels were a little less” (P9).*

Although not all participants found correlations and one mentioned that there was *“no real correlations between stuff I was doing” (P5)*. Although this may be true, it may be beneficial to raise correlations to the attention of users who may not be able to determine them for themselves, or for verifying that there really is no correlation. Another participant suggested that they would like to *“see how your temporal thing is getting correlated with your mood” (P8)*, suggesting that users are also interested in correlations outside of the standard, variable A versus variable B.

## **Determining Factors**

‘Determining factors’ encompasses times participants did not specifically mention looking for comparisons or correlations but rather expressed wanting to determine or find instances in which something happened or could be improved. For example, P9 wished to determine factors that influence happiness and stress, *“I would like to know what kind of solutions, or what kind of factors, I could adapt just to change my habits to become a little bit more happier or dealing with stress in general” (P9)*. P6 also described this as ‘seeing factors’, *“I can see what the factors that affect my studies” (P6).*

Other examples of this using slightly different wording included using recorded data to find calming or comfortable situations, *“to figure out what might calm you down or what situations you feel comfortable in”* (P2) or finding times of best productivity was mentioned by P3, *“to find the optimal times of productivity”* (P3).

The lack of specific analysis to perform might suggest that users may not know how to would go about determining these ‘factors’ but rather would like these ‘factors’ to be drawn out of the data for them.

## Goal Setting

Largely participants in this study did not set goals, this likely stems from participants recording their data in a more *documentary* style rather than *diagnostic* style of tracking. Participants expressed this as not having a goal in mind, *“I didn’t so much have a goal for those when I was doing my task . . . I used them more to see if those tasks would elicit those things in me”* (P7). One participant stated that *“it felt a bit more like a relaxed game so I didn’t feel like I had a goal to achieve just to see how I was performing in a task”* (P9).

Other participants, however, mentioned attempting to achieve specific goals, such as *“to see when I was able to hit high scores in the meditative metric”* (P1). However, goals or target setting was part of three of the metrics recorded from the MindWave and in response to this participants found that the targets were subjective, stating that they were *“a bit confused by how much is high is high, or how much low is low because it was subjective probably it was for someone who’s very very focused mood is high level of mood”* (P8) and thus didn’t know what target to set.

Goal setting therefore presents a challenge when the amount of information available to the user is lacking, making it difficult to determine what goals or targets are appropriate.

## Trends

Participants suggested wanting to find trends over time in their data such as *“different scores across a week and how trends match up over time.”* (P1). One participant highlighted that with the Fitbit’s heart rate there *“was really obvious trends in that and reflected how I felt at the time”* (P2). Previous work has suggested that users use trends as a way to see if they are on track to reach certain goals (I. Li et al., 2011). As such providing users with a means of determining if they are moving in the right direction toward a goal would be beneficial, and there appears to be some overlap



in terms of those participants who mentioned goal setting and those that mentioned wanting to determine trends.

## Summaries

Participants expressed a desire for data to be summarised in different formats. For example, participants felt that the way Fitbit and Spire summarised the data was better than the way Cognizance Tracker presented the data, *“What the Fitbit and the Spire did was they both sent me a weekly summary which was essentially a digest of the data that had been collected.”* (P4). P4 felt that with this type of summary they *“would have started seeing some correlation between what I was doing and what was being observed”* (P4). P1 also expressed a desire for a weekly overview, *“I don’t know if I missed it or something but it would quite good to look at a week’s overview”* (P1). P7 expressed a desire for a summary due to being unable to make sense of the presented graphs,

*I didn’t like that they sort of peak and trough, it didn’t seem like I could really read anything from it because it was just so erratic so I probably would like a summary of that in a way more than just it as a thing.* (P7)

This suggests that some users may not necessarily be interested in the minutiae of the data but rather would prefer a summarised view, sent at various intervals, that they can more easily digest.

In addition to visual summaries, participants expressed a desire for summaries to be expressed in natural language sentences, *“From a layman’s perspective, it can break it down but it can just say that, “OK you’re being this much physical active”, “You’re being mentally occupied at the moment”, those would be nice questions to be answered”* (P8). These comments further suggest that users want the onus of analysing the data to be placed on the PI system, rather than the user themselves attempting to analyse the data.

### 4.3.4 What are the implications of providing people with metrics which have not yet been clinically validated?

The implications of providing users with metrics which are not clinically validated revolve around two central, and related themes, of *‘Trust’* and *‘Accuracy’* of the presented data.

#### 4.3.4.1 Trust

During the post-study interviews participants were asked if they trusted the data they were shown, as a means of understanding what factors may be encountered when users are analysing with data from novel sensor technologies. The responses suggest that participants mostly trusted the values they were presented with. About half of the participants stated that they trusted the EEG values that they were shown and two participants suggested they did not trust the EEG values. One participant said they trusted both the heart rate and breathing rate data equally. One participants stated that they trusted all but the Spire.

A participant suggested they trusted the data despite having no way to tell if the data was wrong, *“I can’t see why in the system it would have either produced wrong results or fake results but then similarly I didn’t understand the reasoning behind the data”* (P7). Whilst another participant suggested that *“there’s a question of what do you believe about whatever analysis it puts forward to you”* (P4).

A factor that appears to influence participants trust included being able to understand how difficult it may be to measure certain metrics,

*I was a bit sceptical about Creativity and Familiarity, was another of those that’s just my kind of assumption, that they would be harder to model and detect I suppose. So yeah, probably just the ones that I felt... I think Attention and Meditation I could trust more.* (P1)

Another factor appears to be an inherent trust in technology lending itself to trusting the data produced, *“I can’t tell you why I did, just an inherent trust in technology I guess”* (P10).

Participant suggested that trust could be gained by showing users how to ‘control’ or ‘access’ wavebands,

*Having similar kinds of games [to Spire], if you like, to explore how one might be able to control at least some, I don’t know whether all would even necessarily be possible, but to be able to see whether conscious control of some of those lines in the Cognizance Tracker were possible. Then I think that would lead to ... being able to have greater trust in the device doing something plausible.* (P4)

When wearing the Spire stone it is possible to ‘control’ your breathing and verify in the companion app that the reading being received is representative of your own breathing

which could act as a means of confirmation. When this is also taken into account with peoples' thoughts on 'live views' of data impacting recorded values (see Section 4.3.2.4) and the likely understandable impact of not attending to the task that is being assessed, it might be difficult to accurately calibrate this trust with sensing technologies that are sensing more abstract information such as the EEG metrics.

#### 4.3.4.2 Accuracy

Participants felt that data accuracy would affect how they would use a novel sensing technology. P1 stated that they would want the data *"to be reflective, if I was to reflect on it and make changes based on what I was telling me."* (P1), suggesting that accuracy is a key factor in how user might reflect on their data. Similarly, P7 suggested that they would want the data to be accurate if they were to continue using it, but stated that *"at the moment it seemed accurate, so if I was going to use it, I would just continue using it in the way that I've said probably"* (P7), suggesting that at the moment that they had no reason to believe any of the data was inaccurate.

P9, however, felt that *"if it was less accurate I would not use it very often. I would just probably rely on the Spire and the Fitbit app individually since it is more accurate I could definitely see myself using it."* (P9), suggesting that users may rely more on sources of data that they believe to be more accurate than other sources of data. An issue related to the previously mentioned tendency to defer to particular sources of data (see Section 4.3.2.5).

P3 felt that accuracy was not as important as being able to discern differences between recorded data (e.g. within a session) for comparison,

*[I]t's not necessary to be very accurate because I've just said that, you don't think about that this is that exact time and then you do this kind of thing. So if there was a half an hour session you are more or less aware for the first part you did this, for the second you did that - so it corresponds to that.*

Participants also felt that increased accuracy would reduce the amount of work that users would have to do for themselves, *"there wouldn't be too much of you trying to work out what was going on"* (P10), suggesting that with less accurately measured data users may have to try and take that into account when performing analyses.

Accuracy has been seen as a challenge of data collection, highlighting that system-driven data collection may result in inaccuracies related to data collection but with

the trade-off that there is less burden on the user to track (I. Li et al., 2010). However, participants responses here suggest that users still expect sensing technologies to provide as accurate a recording as possible. It may be that users are not aware that this is a trade-off they are making.

#### **4.3.5 Does the use of abstracted metric values rather than wavebands have an effect on self-understanding/self-reflection/self-improvement?**

In this section we report the themes of *Difficulty Interpreting Metrics* and *Data Meeting Expectations*. These themes represent the difficulties that were still present when participants attempted to reflect on the abstracted EEG metrics provided in the Cognizance Tracker application, and the challenges encountered when the presented data did not meet the users' expectations of how the data should behave, respectively.

##### **4.3.5.1 Difficulty Interpreting Metrics**

In comparison to the 'obvious' way to interpret step count, participants found that the EEG metrics were still difficult to interpret when presented as abstracted metric values rather than wavebands, *"Although sort of maybe not as intuitive as say like a Fitbit - you want to know your steps that's pretty obvious to anyone what that means"* (P2).

Participants appeared to struggle to understand the values presented as line graphs over time, *"I could see that there was a group of lines that go along at one level and then there's another group of lines that go along at a higher level and that was it."* (P4). Rather than it being an issue interpreting the metric itself, the participant struggled to interpret what these visualised data actually meant, suggesting that appropriate visualisation techniques are needed to help with interpretation.

P7 suggest that both the information about the EEG data could be improved, *"the EEG information could be a lot better"* (P7), and that the way the data is presented could be improved, *"you can see the peaks and troughs but it doesn't mean as much and I would probably rather it be collated together some really easy to understand"* (P7). These comments suggest, again, that further information about the values being presented and the way they are visualised play a role in making presented values easier to interpret. This is likely not specific to the abstract nature of the values themselves but a more general feeling of not having sufficient information and the visualisations not being appropriate for the user.

#### 4.3.5.2 Data Meeting Expectations

Participants suggested that there were discrepancies between how the data was presented and how they felt themselves, *“there was a bit of an internal debate on how I thought I felt versus what it was telling me”* (P2). Despite this, P2 later stated, *“What I would have expected to see, it was pretty much telling me I think”* (P2). Suggesting there may be some internal debate happening between users’ willingness to believe that the data is accurate and trustworthy despite it not matching their expectation.

P3 felt that the values did not match up to their own recorded truth about how they felt, *“I had my mood log and I said I was irritated, angry or tired but from the data it was that I was the most productive or I enjoyed it even.”* (P3). P7 also found similarly that metrics did not behave in the way that they expected in particular circumstances,

*When I was playing games my Appreciation wasn’t very high but I thought my Appreciation would be high because I was enjoying the game and my Meditation was very high which didn’t seem right because I was playing a game which sort of took a lot of thought and processing. So that was a bit weird.* (P7)

Participants also felt that the values did not act in a way that they expected, *“I did find it very up and down on the graph... Which kind of surprised me. I thought it would be more of a smoother trend and it was very up and down.”* (P2)

These comments appear to suggest that the use of abstracted metrics still present their own issues, rather than being a simpler means of presenting data. Despite appearing that participants largely trusted the data they were presented with (see Section 4.3.4.1), there is an issue when the data does not meet their expectations and causes a discrepancy with their own self-evaluation or self-reported truth.

## 4.4 Discussion

### 4.4.1 Discreet & Unobtrusive to Everyday Life

Unlike the automated continuous tracking currently available from wristband or smart-watches, using the NeuroSky headset meant that participants needed to set aside time to track the data for tasks that were of interest to them. Participants therefore described having to setup routines, be consistent in their tracking and that it could feel artificial. The burden associated with manually capturing self-tracking data has been noted by previously by Choe, Lee, Kay, Pratt, and Kientz (2015), suggesting that the

burden associated with manual tracking can compromise the quality of data collected, in addition the burden may negate a technology’s intended therapeutic effect.

During the qualitative analysis of this study we found several barriers to tracking using the provided novel PI systems. The NeuroSky headset design influenced participants tracking practices because it was felt that it was not something that could be worn all day or in all the situations that they would have liked to have recorded data for. Additionally, participants found that that they had to consider things before recording with the devices such as removing clothes, for example, otherwise their recording may be interrupted. For PI systems that require the use of wearables it is likely the ‘wearability’ and ‘unobtrusiveness’ of the hardware therefore plays a key factor, and the burden of tracking with a wearable should be considered. Chuah et al. (2016) found that adoption of smart watches is positively correlated with visibility of the smartwatch, which may be a factor that could be explored further in relation to novel sensing technologies specifically, given that this would appear to be contradictory to comments made by participants in this study.

#### **4.4.2 Knowledge Transfer**

Participants highlighted that they felt they lacked specific knowledge required to interpret their data. Although descriptions of the metrics were provided in the application, participants still felt that they lacked prerequisite knowledge to interpret and make sense of the data. Previous work has highlighted the importance of, and somewhat prevalent, engagement of domain experts in research studies involving PI or self-tracking (Epstein et al., 2020). Ding et al. (2021), in their study of stress tracking with smartwatches, has suggested that we can no longer assume that users of self-tracking tools have the requisite knowledge to understand how data is measured and how to interpret the data being collected. We suggest that designers may wish to include means of transferring knowledge from domain experts to users of novel sensing technology to provide them with the required knowledge to interpret the data.

#### **4.4.3 Simplify Insight Extraction**

Participants struggled to make sense of the data they were provided with and suggested that less onus be placed on the user to have to interpret the data. This perhaps links towards the type of analysis that they wish to perform, or how they wish to understand their data. The abstracted metrics may have provided too low-level data from which to make sense and the provision of automated analyses with less requirement to perform deep analysis of this low-level data would be more beneficial. Another consideration

is that there may have simply been too much information presented for users to make sense of. Similar issues related to too much information being presented to users of PI systems has been noted previously, leading users to struggle to make sense of the presented data due to cognitive overload (Katz et al., 2018). Based on the examples of the types of analysis that users wish to perform from this research and previous PI research which highlights analyses that people use to understand themselves (Choe et al., 2014), it may be possible to perform these analyses on a user's behalf. We further explore whether the use of automatically generated analyses are beneficial to users in Chapter 6. In particular we will explore the value of providing users with both predefined analyses that they can perform, as well as providing automatically generated insights using natural language sentences as suggested by participants in this study.

#### **4.4.4 Build Trust in Novel Sensing Technologies**

The novelty of sensing technologies appears to play a role in encouraging use. However, novel sensing technologies also present concerns about how the data is being handled. Use of the BCI headset highlights that users may consider the data it collects as too personal for the company who is collecting the data or that their data could be used for nefarious purposes. Manufacturers therefore play an important role in ensuring safety of consumer data and ensuring that consumers are aware of the ways in which their data is, or may be, used.

Participants in this study appear to have trusted the data that was presented to them. Although this trust appears to have tendency to be displaced when the presented data does not match users' prior expectations about how the data should look. We further explore how participants determine if a device is trustworthy in the the following chapter. We suggest that blind trust in the data they are being presented with is unlikely to be beneficial, in particular when the devices have not being clinically validated or shown to have reliable levels of accuracy.

#### **4.4.5 Assessing Accuracy in Novel Sensing Technologies**

Related to the previous challenge around trust, participants in this study have highlighted that improved accuracy would be of interest to them and would give them some assurance that what is being recorded can be trusted. However, generally, commercially available sensing technologies, in particular those aimed at the consumer market, do not have regulatory approval and thus cannot make claims about the accuracy or reliability of the data produced. We therefore suggest that designers of novel sensing technologies include a means for self-evaluation of the accuracy of their sensing technologies. Using

a heart rate monitor as an example, this may be as simple as providing instructions to users as to how they can manually measure their pulse using their fingers and a watch. More abstract metrics and things that are more difficult to accurately self-measure require further research to determine the ways in which this can be achieved without impacting the recording of the task the device is intending to measure.

## 4.5 Limitations

Participants were self-selected and thus are likely representative of a population that are open to taking up novel technologies and not necessarily representative of the general population. Additionally, as participants in this study were largely recruited from the University, the results may not generalise to other populations of users.

That majority of participants did not have a particular goal in mind and are likely to fit into the documentary and/or fetishized style of self-tracking (Rooksby et al., 2014). Therefore directive, diagnostic and reward orientated self-trackers may approach understanding novel sensor data differently. We further explore the challenges of novel sensing technologies used for diagnostic or directive tracking in Chapter 5.

## 4.6 Conclusion

In this chapter we presented the results of a qualitative analysis of 10 participant interviews conducted after a three-week in-the-wild study involving a multifaceted PI system capable of recording and presenting data captured from three sensing technologies: Fitbit heart rate, Spire breathing rate, and NeuroSky EEG metrics. We highlight that the unobtrusiveness of wearables likely has a role to play in uptake. Additionally, we suggest that means of transferring knowledge and providing means of simplifying insight extraction are likely to be beneficial to users.

We suggest that the design challenges and challenges presented in this chapter are likely generalisable across a range of tracking styles given that they do not appear grounded in a specific style of tracking but rather are challenges based around use of novel sensing technologies more generally. We explore challenges related to directive and diagnostic tracking styles in the following chapter. We discussed challenges related to trust, both from the perspective of users' data and from the perspective of users trusting the data that they are presented with. Based on our findings that users tended to trust the data they were presented with and lacked means of assessing the accuracy of data from the provided sensing technologies.



Issues related to *Trust* are explored further in Chapter 5 and Chapter 7. The challenge of *Insight Extraction* is further explored in Chapter 6.

## Chapter 5

# Making Sense of Novel Sensor Data Within The Context of Fatigue

## 5.1 Chapter Overview

One limitation of the study described in the previous chapter was that it did not include people who feature in the diagnostic and directive categories of tracking, and how those people might use a novel and multifaceted PI system. People who track data for diagnostic or directive reasons are described by Rooksby et al. (2014). People who track data in this way are tracking to understand or determine if there are things that can be changed, which may be beneficial, such as improved quality of life. For example, they may seek a better understanding of their illness or a better understanding of themselves, or how to best manage their symptoms. The study described in this chapter aims to provide a further understanding of how those who do track for diagnostic or directive reasons do so and the challenges they face.

In this chapter we described a study in which participants made use of a mobile PI tools capable of recording data from a NeuroSky MindWave, Fitbit and Spire, over a three-week period. This follows a similar format to the study described in the previous chapter. However, in this study we sought participants that are living with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS), prolonged fatigue, chronic fatigue (CF), and idiopathic chronic fatigue (ICF). ME/CFS is a long-term multi-system illness, the symptoms of which include post-exertional malaise, impairment of physical or cognitive activities that is not improved by resting, and unrefreshing sleep (Clayton, 2015). People diagnosed with ICF may also experience post-exertional malaise but without symptoms that impair physical activity. Chronic fatigue is defined as fatigue lasting longer than six months. Prolonged fatigue is self-reported fatigue lasting longer than one month but less than 6 months. An algorithm for differential diagnoses of these conditions can be seen in Figure 5-1. People living with these conditions have a variety of symptoms and the conditions are not interrelated. However, they all share one commonality, related to varying levels of ‘fatigue’ or malaise.

Currently there is no automated objective way of measuring a person’s fatigue level and measurements are usually based on subjective ratings, multiple question surveys which aim to provide an outcome measure that can be used to determine peoples’ levels of fatigue, or performing tasks to determine fatigue levels (Brunet, Dagenais, Therrien, Gartenberg, & Forest, 2017). There is current research that suggests that there may be correlates in EEG data with peoples’ mental fatigue levels (G. Li et al., 2020). Levels of fatigue/attention extracted from EEG data are already being explored by researchers in the context of critical situations such as driving (Gharagozlou et al., 2015) and surgery (Ndaro & Wang, 2018). Additionally, research has sought to understand the

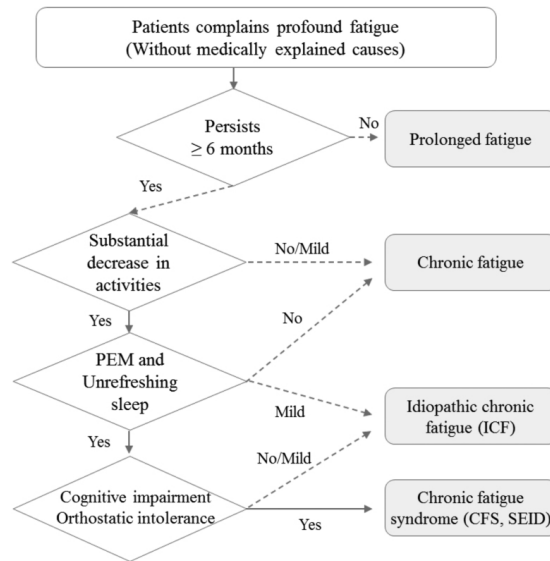


Figure 5-1: Suggested diagnosis algorithm, showing differences between different fatigue presentations (Son, 2019)

changes in EEG data in those living with ME/CFS (Zinn, Zinn, & Jason, 2016).

In the context of PI, Davies et al. (2019) found that people living with ME/CFS attempt to use commercially available technologies, including wearables, and are interested in determining factors that effect their fatigue levels. Research has suggested that there may be benefits from attending to heart rate data to try and manage fatigue or malaise in those living with ME/CFS (Windthorst et al., 2017). This type of tracking exemplifies the diagnostic and/or directive styles of tracking.

The main research questions that we intend to answer in this chapter are:

**RQ1** How do people with ME/CFS perceive benefits and limitations of collecting data from novel sensing technologies?

**RQ2** How do participants attempt to make sense of data from a range of novel sensors?

Based on the issue of trust highlighted by the previous study we also attempted to elicit further information around trust in PI systems to understand:

**RQ3** How do people determine if a device is trustworthy?

## 5.2 Methodology

To better understand how a particular population of people might use these technologies for directive or diagnostic tracking purposes people living with myalgic encephalomyelitis/chronic fatigue syndrome (ME/CFS), prolonged fatigue, chronic fatigue (CF), idiopathic chronic fatigue (ICF) were provided with an updated version of the Cognizance Tracker application, described in the previous chapter, to use over a 21-day period. Participants in the study were provided with a NeuroSky MindWave, Fitbit Versa and Spire stone, similar to the previous study. Participants were interviewed after 21 days using a semi-structured interviews. A courier was used to deliver hardware and the interviews were conducted over Skype or telephone, to limit travel requirements for participants. This provided an additional benefit of not being restricted to recruit participants who were local to the University of Bath. Participant briefings took place over Skype, and lasted about 30 minutes to an hour depending on the participant. Time was spent with each participant to explain how to use all of the hardware. Before the briefing ended participants had made a short recording session of a few minutes, to ensure they understood how to place the sensors and that they understood the what was displayed during a recording session. Participants were also encouraged to reach out to the researcher if they had any issues during the study period (e.g. for technical assistance). The study was run over a 21-week period.

The study methodology received ethical approval from the university’s Research Ethics Approval Committee for Health (REACH) (Reference: EP 18/19 028). Data pertaining to this study is stored in accordance with the data management plan in Appendix F.

### 5.2.1 Application

Building on the application used in the previous study, this iteration of the application still allowed the recording of NeuroSky eSense values as in the previous version, as well as manually recording mood data using an Affect Grid (Russell et al., 1989). The updated version of the application removed the ability to review data in real-time while the session is being recorded and no longer recorded data using NASA TLX. The live view was removed based on the results of the previous study, where participants suggested that the live view was a distraction from the real task being tracked. The NASA TLX was removed based on a consideration of the types of tasks that participants in this study were likely to undertake and a consideration of the complexity of having to answer the NASA TLX questions. An additional section enabling the user to record fatigue ratings was added, using a 10-point fatigue rating scale (Micklewright,

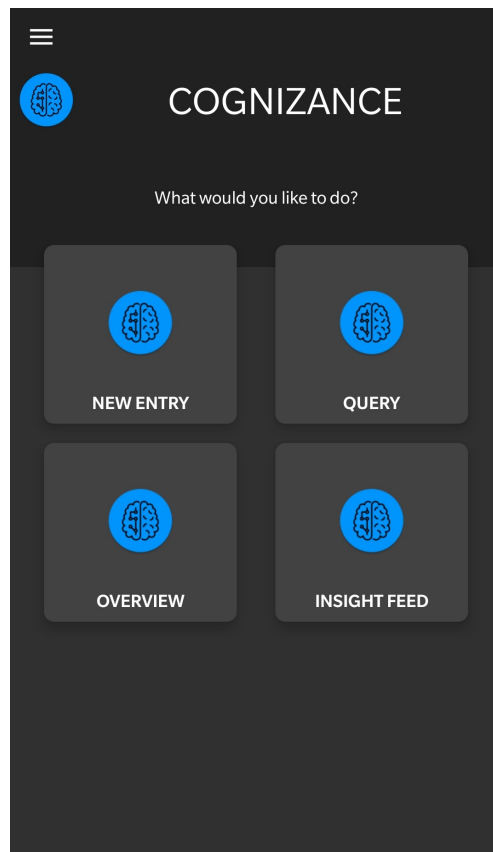


Figure 5-2: Cognizance Tracker Main Menu

St Clair Gibson, Gladwell, & Al Salman, 2017), given the likely benefit of such a value to the intended study participants. The steps of a recording session can be seen in Figure 5-3.

Some UI changes were also made to the session overview such as: changing pie charts to bar charts, adding the ability to hide or show specific chart sections (see Figure 5-4). Additionally, the following changes were implemented to enable easier usage of the application based on suggestions from Davies et al. (2019):

- The minimum font size was increased to 14pt to ease readability.
- The application also included the ability to switch between two colour modes. The alternative colour mode, Dark Mode, was implemented specifically as it was noted by that this would be beneficial to people living with CFS/ME by presenting everything on a dark background (with complimentary foreground colour).

The updated version of the application contained two new features, an Insight Feed

and Query Area (see Figure 5-5). The features were included to investigate whether either would be beneficial in enabling people to explore or gain further insights from their data. Further details and analysis of the Insight Feed and Query Area can be found in Chapter 6.

←

New Recording

⋮

SKIP TO MOOD RATING

CONNECT HEADSET

Signal Quality:

START

SPLIT

STOP

Session length: 00:00:00

←

Mood Rating

⋮

SKIP TO FATIGUE RATING

SAVE

Select your mood on the graph below

stress

high arousal

excitement

unpleasant feelings

pleasant feelings

depression

sleepiness

relaxation

Use this area to record additional details

e.g. What are you doing? Why are you feeling this way?

←

Fatigue Rating

⋮

SKIP

SAVE

Please record your current level of fatigue on a scale of 0 to 10, where 0 is not fatigued at all and 10 is total fatigue/nothing left.

0

5

10

Selected value: 0/10

0

1

2

3

4

5

6

7

8

9

10

Not at all fatigued

A little fatigued

Moderately fatigued

Very fatigued

Total fatigue & exhaustion - nothing left

Use this area to record additional details about your fatigue.

e.g. reason for (not) being fatigued, things to think about in future...

SAVE

Figure 5-3: Steps of New Recording



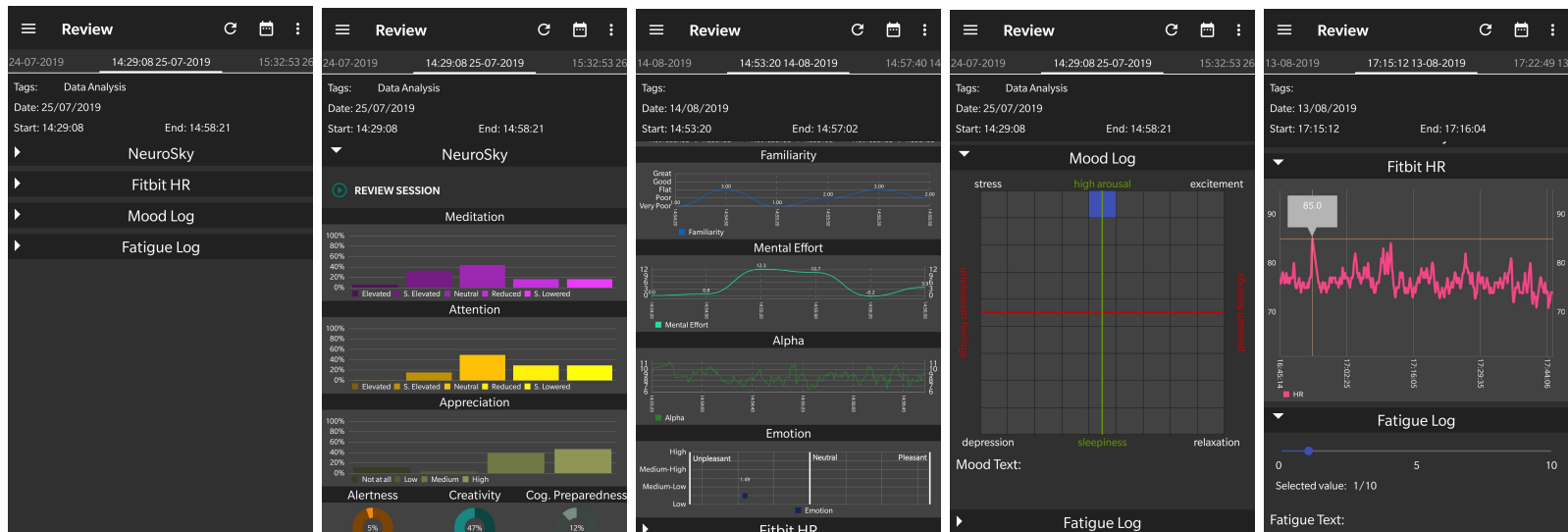


Figure 5-4: Session Overview

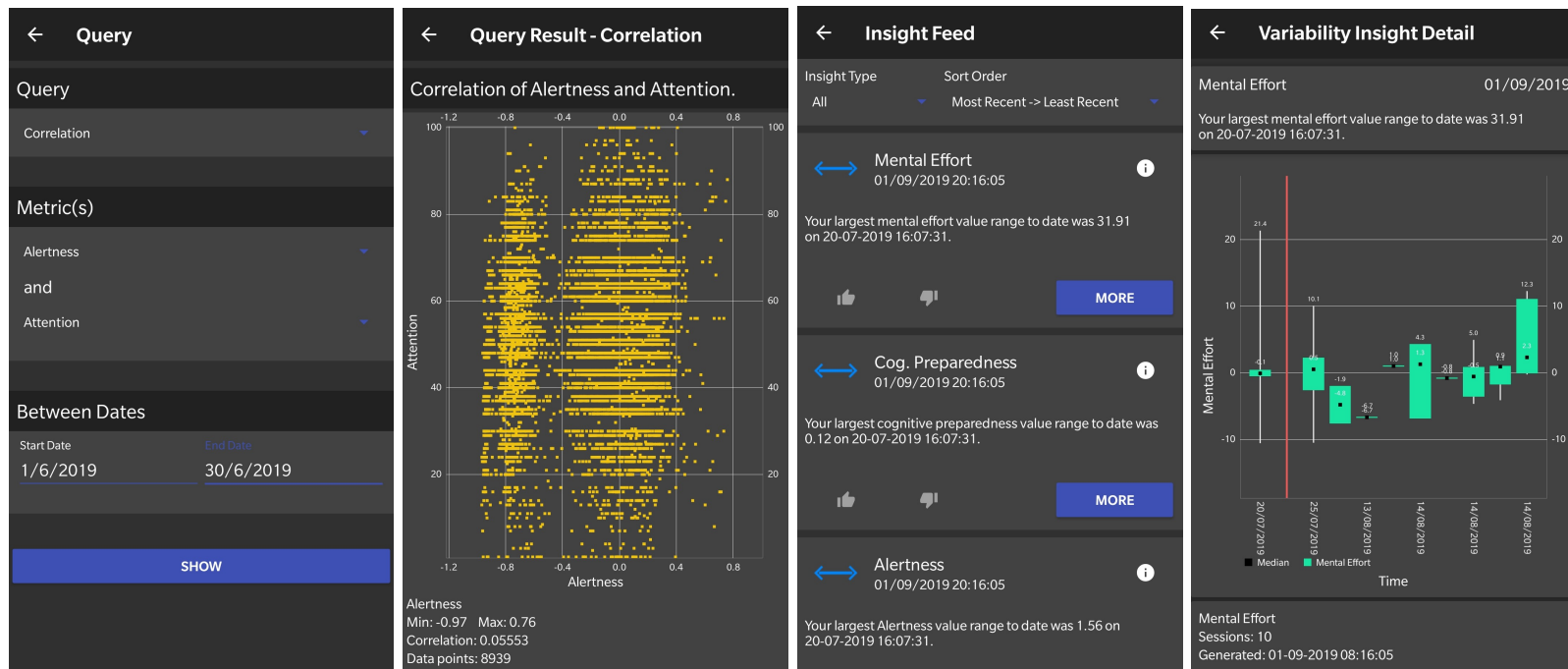


Figure 5-5: Query Area &amp; Insight Feed

### 5.2.2 Participants

Participants were recruited through a variety of channels including Facebook support groups, Twitter posts, and the use of an online participants recruitment platform, Call For Participants<sup>1</sup>.

Nine (M: 4, F: 5) participants took part in the study with ages ranging from 23 to 75 (M: 45.67, SD: 17.69). P6 is intentionally missing from presented results having withdrawn after starting the study. One of the participant interviews conducted had the participant's parent answer in addition to the principal participant as they had been involved in helping to track and monitor the data on their behalf. Of the participants who completed the study, 6 participants are living with ME/CFS, 2 with CF, and 1 with ICF.

Participant #	Age	Gender	Diagnosis
P1	44	M	ME/CFS
P2	75	F	ME/CFS
P3	61	F	ME/CFS
P4	36	F	ICF
P5	34	F	ME/CFS
P7	57	M	CFS
P8	56	M	CF (Cancer)
P9	25	F	ME/CFS
P10	23	M	CF (Ehler-Danlos Syndrome)

Table 5.1: Study 3 - Participant Demographics

Participants were asked to use the provided technologies over a three-week period. Afterwards participants were interviewed to gather feedback about how they used the technology, what they liked, what they wanted to be able to do, if they were able to do that and if not what they would need in future to enable them to do so.

The number of sessions varied per participants and ranged between 7 and 33 ( $M=19.67$ ,  $SD=8.03$ ) sessions over the three-week period, as can be seen in Table 5.2.

Interview results were analysed using thematic analysis Braun and Clarke (2006). The interviews were transcribed in full, a total of 06:27:17 worth of audio was transcribed. These transcripts were then read and re-read to become familiar with the data. The transcripts were then reviewed and coded inductively, using the data to direct the creation of themes. The themes from the study described in Chapter 4 were used as an initial reference, as such there are some themes that overlap. However, additional

<sup>1</sup><https://www.callforparticipants.com/>

P #	Recording Sessions
1	24
2	21
3	24
4	33
5	9
7	14
8	18
9	27
10	7

Table 5.2: Recording Sessions Per Participant

new themes were developed due to the specificity of the responses not fitting within the themes developed in the previous study. The themes were reviewed and are presented in the following sections. An overview of the themes and their related research questions is presented in Table 5.3.

## 5.3 Interview Results

### 5.3.1 How do people with ME/CFS, CF or ICF perceive benefits and limitations of data from novel sensing technologies?

Perceived benefits and limitations of tracking using novel sensing technologies were described in terms of the *Immediacy of Data* being recorded, *Difficulties Due to ME/CFS, CF or ICF*, and *Tracking Data of Interest* to participants.

#### 5.3.1.1 Immediacy of Data

Participants expressed a desire for interventionist technologies, “*knowing that at a particular moment in time where you are level-wise is also quite important*” (P1), similar suggestions were seen in our previous study. Immediacy was suggested as tools that could objectively predict that someone was likely to become fatigued, or were reaching some predetermined threshold, “*So some means of identifying where is the point where you still feel OK but actually your EEG data is telling you you’re getting into the danger zone*” (P3).

Participants also perceived that measuring heart rate using the Fitbit gave a more immediate sign of impact than using the headset. Participants highlighted that in some instances they were checking the Fitbit watch every five minutes to see their status, “*I probably glance at it every five minutes at least just to make sure that my*

Theme	Sub-theme
<b>5.3.1 How do people with ME/CFS, CF or ICF perceive benefits and limitations of data from novel sensing technologies?</b>	
Immediacy of Data	Longer Term Trends & Data Collection
Difficulties Due to ME/CFS, CF or ICF	
Tracking Data of Interest	Tracking Context Data of Interest
<b>5.3.2 How do people living with ME/CFS, CF or ICF attempt to make sense of data from a range of novel sensors, for which they have varying degrees of understanding? What problems do they encounter?</b>	
Target Setting	
Comparisons	Comparisons to Others & Themselves
Data Not Meeting Expectations	
Insufficient Knowledge or Information	Theory & Research Basis
Data Presentation	Amount of Data Present Usefulness of Data Presented Technicality of Data Presented
<b>5.3.3 How do people determine if a device is trustworthy?</b>	
Own Research	
Compare and Correlate Sensor Results	
Inherent Trust in Technology	
Trust Values That Meet Expectations	
Trust Values That Are More Accurate	

Table 5.3: Summary of Themes

*heart rate hasn't suddenly jumped up*" (P1). Participants also highlighted checking the Fitbit app itself two or three times a day, or having built a routine of checking their Fitbit data each morning.

These suggestions possibly highlight that not all users are likely to want to use technologies to record and analyse the data in any deep way. Participants comments suggested that they would prefer to use the technologies for more immediate reflection by checking statuses, which has been documented previously by I. Li et al. (2011). There also appears to be a preference not to use PI tools but would rather there be some sort of automation that would tell users what to do or that they are reaching limits that have been predetermined by a third-party.

### **Longer Term Trends & Data Collection**

Not all participants suggestions were related to immediacy of data but rather they would prefer to track and analyse data over a longer period of time. The changes that participants are interested in are suggested to be long-term trends over much longer periods than three weeks.

Participants suggested a willingness to take part in research studies that were in the region of up to six months, suggesting that three weeks was not long enough to get used to the technologies and to collect enough data to reflect on, *"with three weeks of recordings I didn't have enough chance to contrast those particular sort of things. But that's the sort of thing I might have chosen to look at if it was a six month study"* (P3). P5 highlights, from previous experience, the benefits of longer-term tracking which enabled them to determine that a change in medication had resulted in reduced strain on their body,

*So because it measures the steps as well you've got an objective measure of your physical activity and then if I've done the same number of steps over three months, per month over three months, but my intensity minutes are fluctuating then I can tell that. For example, recently I could tell because I changed my medication that even though my number of steps hadn't changed over three months the number of strain that my body was under when I was doing them had gone down.* (P5)

Further to this, participants suggested that they would likely want to be able to track trends in their data for periods of time such as 18 to 24 months, *"resting heart rate is a long term trend, and I'm talking over eighteen months, two years, is something of interest but it's very difficult to get that"* (P1).

PI tools that allow for tracking of longer term trends have been previously highlighted by Rapp and Cena (2014), suggesting that additional work is required to encourage engagement with tracking longer-term trends. Although this may be an issue, it would appear that at least some users are willing to track longer periods of their own accord, based on their own understanding of the necessity to sustain tracking for longer in order to gather the required information for PI tools to be useful to them.

### 5.3.1.2 Difficulties Due to ME/CFS, CF or ICF

Participants expressed issues analysing data due to issue related to their condition. In particular participants suggested that they often found coming to terms with a new technology or a device, fatiguing in and of itself, *“Obviously again it’s down to the fatigue itself. I find learning things takes a lot longer”* (P7).

Participants also expressed that the act of analysing vast amounts of data could be fatiguing on its own, *“one of the issues that’s really common is brain fog and it’s kind of how it’s described and it just means that sometimes having a lot of data in front of me can be quite difficult”* (P9), *“being fatigued means accessing something that’s complicated and new is actually really fatiguing”* (P8).

P3 noted that the act of recording and then attempting to analyse data immediately afterwards also contributes to fatigue,

*That might be partly ‘ME-brain’ because that sort of thing is particularly challenging. When you’ve doing whatever you’ve done for the session and then immediately go in and start looking at it you start getting into overload and then you’ve got to try and remember when you were when you come back to look at it, or come back and look at the insights the next day when they come.* (P3)

One participant, in describing the amount of information shown in the application, suggested that it should be “streamlined” to help navigate the data more easily, reminding them of events, such as when they were newly diagnosed and being unable to read for 3 months,

*I think my main thing would just be streamlining it because, you know, one of the biggest difficulties, one of the most widespread difficulties of people with chronic fatigue is navigating data, like, you know, for example, when I first got Chronic Fatigue I couldn’t read for three months because I didn’t have the energy to separate out the words. Uh, so I think just it’s just a bit*

*busy at the moment.* (P9)

This highlights an important consideration in determining the best way to allow users to record and analyse data that takes into consideration difficulties the user may have based on their condition. Although the design should take into consideration more factors than simply providing less information, as this could just prevent people accessing data that may be helpful.

This issue was also noted by Davies et al. (2019), which highlighted that excessive amounts of data could trigger symptoms of fatigue. It may have been beneficial therefore to have included participants in the design of apps that they are intended to help. Although we had initially considered running participatory design as pre-work to this study which could have helped design a more tailored application, we were limited by a lack of tools that could support conducting participatory design remotely - a necessary requirement when conducting research with people who may not be able to travel.

#### **5.3.1.3 Tracking Data of Interest**

One significant limitation of any PI tool is not providing data of interest to users. The sub-themes of *Tracking Context* and *Data of Interest* highlight the contexts that participants wanted to track data about and the specific types of data that participants felt were, or would be, useful to better understand themselves.

#### **Tracking Context**

Participants in this study were also asked what they were doing when recording and what other things they would have liked to have made recording sessions about. Table 5.4 shows the occurrences of activities mentioned by participants during the post-study interview. Several of these overlap with contexts participants from the study described in Chapter 4 wanted to record data about. Although *Physical Activity* was mentioned the most, the majority of contexts for comparison are between ‘sedentary’ tasks. Activity trackers tend to focus on more active tasks, however, the contexts in which users may want to track data might not always be what is considered physical activities. However, it is worth noting that in the previous study participants also mentioned numerous sedentary tasks, which may be a result of the headset being perceived as a device which can only be used in such activities.



Context	No. Participants
Physical Activity	5
Watching TV	4
Computer work	4
Reading	3
Gaming	3
Talking on the phone	2
Cooking	2
Chores (e.g. hoovering, washing up)	2
Reviewing bank account	1
Physio	1
Resting	1
Eating	1
Studying	1
Driving	1
Sleeping	1
Gardening	1
Playing Piano	1
Sitting on the sofa	1
Working	1

Table 5.4: Variety of Recording Contexts

## Data of Interest

In addition to contextual information, participants mentioned tracking data from specific sources to better enable them to manage their symptoms. Participants were keen to determine things that may influence their levels of fatigue or energy expenditure. Participants showed interest in manually tracking fatigue levels and tracking their alertness levels, in relation to the app and devices provided. Participants also suggested tracking information such as medications to determine their effect on their symptoms, tracking vitamin levels, balance, and social interactions would be useful to them.

### 5.3.2 How do people living with ME/CFS, CF or ICF attempt to make sense of data from a range of novel sensors, for which they have varying degrees of understanding? What problems do the encounter?

Participants in this study described encountering challenges with *Target Setting* and *Comparisons*. Participants also encountered issues related to *Data Not Meeting Expectations*, *Data Presentation*, *Insufficient Knowledge or Information*.

### 5.3.2.1 Target Setting

Participants mentioned previous experience with target setting using heart rate ranges and finding it beneficial,

*I probably get more benefit from recording heart rate data because the resting heart rate generally gives me a trend that tracks how I'm feeling. So a higher resting heart rate generally my symptoms are worse and I'm feeling worse. Within the day if I'm doing an activity it gives me something that I can sort real-time look at. If I'm doing something in the kitchen, if I'm making a cup of tea and I'm at 120 beats per minute I'll think ah it's one of those days I just need to sit down and rest not do a 10 minute activity like I would done two days before say. (P1)*

Another participant was interested in using Fitbit to stay within specific heart rate zones but found that Fitbit did not allow for this without installing another app. However, using this app prevented the Fitbit from being used for anything other than tracking heart rate,

*Trackers always have heart rate monitors on them, or tend to now, and one thing I'm particularly using is I'm trying to remain below a certain rate all the time, and the only way on the version of the Fitbit you sent me to do that is to use one of the little apps that you can add, and it basically assumes that you're doing exercise so then I can't actually look at the face of the watch, I can only see my heart rate. So I would like the ability to do these things for... You know, just to have these things designed for people who are doing things other than training to get incredibly fit because there are other reasons we want to track things. (P4)*

This highlights a further need for designers of tracker to consider how people outside of those using them for fitness may be limited by such designs.

Participants also expressed concern about target setting due to previous bad experiences,

*Whether it's a good idea, is another matter because I've done previous, having a personal trainer set up a fitness plan and building up over several months and unfortunately I hit a sort of brick wall at one point and was set back quite a long time so I'm wary of doing it. That's not to say I wouldn't do it but I have my, sort of, apprehension about doing it. (P7)*

This highlights a concern, that we raised as part of this study’s information sheet, that users may be inclined to follow the data and end up pushing themselves too far, with unintended negative consequences. This is an important point that should be considered when designer PI tools, in particular for those who use the tool to direct or change their behaviours.

Participants also suggested they lacked information to set target, *“I think again I would need more information to start with to know where I should be trying to set my targets.”* (P3), and that target setting may not be particularly useful in the context of fatigue,

*If I go for a swim, which I do on a weekly basis to try and improve my fitness, I don’t set myself a number of lengths. I don’t count how many lengths that I swim. I just do what I can manage that, you know, I’m doing a little bit would be useful. But setting myself targets isn’t very useful in terms of fatigue, because it isn’t a gradual build up like an athlete would do in performance. It’s, you know, take it, take it each day as it as it comes.* (P8)

These comments suggest that target setting may be quite individualistic, and that not all users will be keen to make use of them. Dissemination of information both in terms of what targets could be set and how target setting might be useful to users could be beneficial.

### 5.3.2.2 Comparisons

In much the same way that participants in the previous study suggested that they would like to compare metrics between values, participants in this study also wished to be able to do that but encountered specific issues in this regard.

#### Comparison to Others & Themselves

Participants suggested that they lacked prior data to be able to compare data to themselves pre-diagnosis, *“I had no idea, if I went back pre-ME, whether they would have looked any different or not”* (P3). Participants highlighted being able to see changes in the data but not having the required understanding, information or alternative data to compare it to,

*So I didn’t have anything to put it against apart from another reading. I found with the measurements, for instance thinking about the Alpha meas-*

*urements you could see whether it had gone up and down but why I didn't know was whether it was too high or too low and how that related back to me. (P3)*

In addition to this participants expressed missing contextual information that was not present in the app, by not knowing what ‘normal’ values would be, or what ranges would be considered optimal,

*I would love to know what it's actually showing in terms of difference between me, with Chronic Fatigue, and somebody without Chronic Fatigue. I'd be really interested to know what difference that proves to be. (P7)*

Benchmarks were suggested by one participant, but highlighted that due to the individual nature of CFS it may not be suitable to compare to other people with CFS. Their comment highlights a need to find an individualistic means of providing users with relevant information to perform comparisons with (if deemed appropriate),

*So benchmarks, and that would include benchmarks of the average person. It would be good to split it by gender. It would be good to split it by age. For Chronic Fatigue it's not necessarily helpful to compare to other people with Chronic Fatigue because it's so variable for people. (P4)*

### **5.3.2.3 Data Not Meeting Expectations**

Participants expressed that the data they reviewed didn't behave the way they expected. P1 felt that the data behaved more noisily than expected, *“There was so much variation rather than it being a . . . I was going to say smooth but not a smooth trend but I was expecting something that was lower frequency in terms of movement.” (P1).*

Participants also struggled to correlate how the data was graphed with how they felt in the moment, *“I think there could be benefit in it but the thing that I struggled with was trying to correlate what the app was saying against my experience.” (P1)* and further elaborated by P9, *“I might feel very cognitively unprepared and very unalert but the device was saying that I was but I'm not saying that that device means that it was wrong. I just I think it's just not understood” (P9).*

Participant felt it was difficult to understand how the values presented could be interpreted in combination with other values, *“I was a bit mystified about why a large Mental Effort didn't appear to go with a large amount of Attention and Alertness and Cognitive Preparedness” (P3).*

Participants also suggested that the data was contrary to what they expected. For example, participants mentioned seeing high levels of alertness whilst they were asleep, *“I seemed to be more alert when I was asleep than when I was driving for half an hour, which is a puzzle to me, because I think I was fairly alert when I was driving”* (P8), or seeing high levels of unpleasant emotions when watching TV, an activity which they enjoyed, *“I was also quite intrigued I often got more unpleasant emotional readings that I expected on these. So for example it told me that I was in the less pleasant emotional reading when I was watching the rugby”* (P3).

Participants comments also suggest that data that does not behave in the expected way would lead them to question the data provided, *“[I]t recognised I was not at all appreciative. Probably if it had said very appreciative, I would have a very considerably questioned the definition they were using of appreciation”* (P10).

Users would likely gain benefit from PI tools that manage expectations of what recorded data might look like. This is perhaps also somewhat related to the issue of target setting and comparisons, in that although users may be recording and analysing data about themselves, first there must be some guidance about what one can expect from their data to be able to begin making sense of it on a personal level.

#### **5.3.2.4 Insufficient Knowledge or Information**

Again during this study participants expressed lack of sufficient information or missing knowledge as a factor in attempting to make sense of the presented data. Some of which has already been noted in regards to *Target Setting* and *Comparisons*, however, this theme focuses on lacking information in relation to the metrics, data being presented, or sensing device itself. As was also seen in the previous study, participants felt they required specific neuroscience knowledge to better understand their data,

*I think it's potentially useful but me not being someone who knows much about, kind of, neuroscience and the value of EEG measurements and the accuracy of the device itself I found it was a bit difficult for me to really know whether it was valuable or not.* (P5)

Participants who attempted to find out further information about the NeuroSky headset felt that the company's website was aimed more so at developers and did not provide accessible information that they could make sense of,

*I tried to look on the website of the actually, kind of, MindWave, kind of, information on their website and I didn't find it very easy to find information*

*that kind of told me, “OK this is what it is, this is how it works” a lot of it seemed to be aimed at developers, not necessarily aimed at a lay user to explain. (P5)*

P3 felt that “*something like Creativity, without knowing what is it using to pick up that information, how does it come up with the fact that was either 0 or 50 percent or 100 percent?*” (P3) more difficult to interpret, affecting their trust in that specific measurement. This suggests that greater information about how the sensors work and how specific values are determined might have helped them make better sense of their data.

## **Theory & Research Basis**

Participants expressed a concern that the values provided by the EEG sensor may not be grounded in theory or prior research. P10 from their own research said,

*I recently read Stephen Jay Gould’s book ‘The Mismeasurement of Man’ and it’s made me very leery of proxy matches, saying that when you’re brainwaves make this pattern, it’s to do with this kind of thought. To me it feels like saying doing good on IQ tests means that IQ is a real thing that determines mental ability and human worth. It sort of feels like a big jump that needs heavy justification and none of the materials provided gave that justification. (P10)*

Some participants appear to be aware of current research about what things might be helpful in managing their condition, such as tracking of HRV as a surrogate marker for energy expenditure,

*“They do a lot exercise physiology studies and they use heart rate within that and then there are, kind of, management strategies that some of the researchers advise people to use based on heart rate being a rough measure of anaerobic threshold and things.” (P5)*

When asked about the benefits of recording EEG versus heart rate, P5 felt that they were more inclined to find heart rate feedback more actionable, suggesting that “*a big part of it is that there’s a theory behind it and it’s an established part of the literature on the condition anyway. So it’s much easier to interpret whereas if you’re just interpreting data without that framework it’s hard to make meaning out of it I think.*” (P5).

As such participants appear keen to be able to refer to and track data which is grounded in prior research, both as a means of knowing that the technologies are validated and

are useful but also that they could refer to and understand what they should be tracking and how they should be tracking it. This conflicts with our earlier results (Chapter 3), which seemed to suggest that users were not concerned about the scientific basis on which the system provided information. However, although prior study participants may have showed less interest in this, the benefit of providing this information is still likely to be of benefit to them as previously suggested.

### 5.3.2.5 Data Presentation

Issues related to data presentation fall into three sub-themes: *Amount of Data Presented*, *Usefulness of Data Presented*, and *Technicality of Data Presented*.

#### Amount of Data Presented

Participants suggested that there was too much information shown in the application. This is perhaps somewhat related to their condition (see Section 5.3.1.2). However, it is worth noting that this was also an issue noted in the previous study and thus copious amounts of data is unlikely to be seen as good.

Participants felt that they were having to filter the data themselves to try and make sense of it, *“it’s almost like as the user you have to apply your own filter over the top to try and interpret it”* (P1).

Another issue highlighted was that the form factor, of a mobile phone, was not supportive in analysing the amount of data shown. Participants suggested that they would prefer to analyse data on a large screen such as a tablet or desktop computer. P3 would have preferred a larger screen to perform analyses, *“it would have been good that have been able to download that somewhere onto the computer and then look at it on a bigger screen.”* (P3), suggesting that although they liked the graphs could be zoomed in it still didn’t provide the size preferred, *“I like the fact you could expand the graph on screen although I still found it quite small.”* (P3). P10 echoed this sentiment stating,

*I think to use it on the on a tablet would have been more comfortable than using it on the phone. But to have the web application as well, so that you can compare it data in different ways. Um, you know, compare the data from two or three different sessions at the same time on a large screen, a laptop or a desktop would have been very useful.* (P10)

Larger real estate to analyse data on is often provided, such as having a larger dashboard on the manufacturer’s website (e.g. Fitbit). With more and more sensors being built

into mobiles, wearables tending to synchronise data with mobile devices, and with increasing trends towards ‘mobile first’ web design and mobile-only apps, it remains important to understand user preferences in terms of being able to analyse their data with these responses suggesting that users still want to make use of larger displays when analysing their data.

### Usefulness of Data Presented

Participants also felt that some of the presented data was not useful. This may be somewhat related to the amount of information shown and a need to provide users with a means of filtering data that is of interest, *“There was only certain things I was interested in looking out... And I feel like there was quite a lot of different information.”* (P1). As the metrics provided from the NeuroSky headset were not filtered to be what we considered relevant for participants this is perhaps somewhat to be expected that not all metrics were perceived as being useful.

P10 felt that the usefulness of the data presentation was very important, particularly if they were gathering more data, *“I don’t know whether gathering more data will be useful if it is not being presented usefully, or if the analysis is based on assumptions”* (P10).

Participants highlighted that, for example, sleep data as provided by Fitbit was measuring something ‘tangible’ whereas Creativity, as measured by the NeuroSky, was not tangible,

*When it’s telling you how you slept last night. That’s very tangible. It’s how restless and restful, whether you’re awake or not. That’s useful data but saying how creative, according to you... Saying how creative you are based on the fact that we told some people to be creative and measured the squiggles, or asked people to wear this stuff and then asked them when they were being creative and assumed that the squiggles while they said they were being creative, actually correlates with creativity. It’s sort of... I know when I’m being creative, and that’s when I’m seriously typing away, trying to capture an idea before I forget it.* (P10)

One participant suggested that if the metrics could be combined into a ‘super metric’ this would be more useful, *“I wonder whether the nature of what it is you’re recording it doesn’t lend itself to that unless you can come up with some snazzy way of crunching all the data and coming up with some other super metric that you combine as the inputs”* (P1). Suggesting that some users may prefer to get answers from a minimalist presentation of their data rather than being provided with as much information as pos-



sible. Alternatively, participants suggested that metrics could be presented as colours representing, for example, correlations rather than values, *“I don’t feel like I would need a specific data output, I would just need to know generally what was happening. So whether there was a correlation, it could be green. Or whether there wasn’t it was red”* (P9). Suggesting that statistical values are not useful to all users and they would prefer simpler designs to convey the information.

Deterministic or diagnostic styles of tracking applications may find it helpful for users to be able to turn on or off specific sources of data rather than providing all available metrics, or sources of data, at once. However, it is possible that users would choose to only include sources of data that provides what they consider to be useful information, either by only buying sensors which provide that data or only importing data of interest in the case of a multifaceted PI tool. Additionally, users may find simple UI designs that are not based on data-heavy UI elements more useful.

### Technicality of Data Presented

Another data-related issue was that participants felt the data presented was too technical, rather than user-friendly. Participants would have preferred that data be presented in a less technical manner,

*I read as much as I could of the information that was on the app and uh, you know to try and understand what the different areas were about. What Alpha waves are, what valance means, all that kind of thing. Um, but I was a bit, I’m a bit confused. I guess I’m a little confused about it, it seems quite technical rather than user-friendly I guess.* (P8)

P4 echoed this sentiment, suggesting that the values would be better presented in ‘layman’s terms’, *“Putting them into layman’s terms. So something the average person in the population is here and you are here, rather than just giving a number”* (P4).

Katz et al. (2018) usage study of how people living with Type 1 Diabetes interact with self-tracking apps, seems to present similar results whereby users can be overwhelmed by the amount or content of the data presented, or issues related to lacking specific statistical knowledge, or specific statistical knowledge leading to questioning the utility of the presented data. Therefore this may not be an issue specific to just one set of potential users and should be considered more generally when designing PI apps.

### 5.3.3 How do people determine if a device is trustworthy?

Continuing the thread related to trust from our previous studies, we again asked participants whether they trusted the values they were shown and we asked those who had previously used tracking technologies how they decided if a device was trustworthy. The ways in which participants described how they determined trust are described in the themes of *Own Research*, *Inherent Trust in Technology*, *Compare and Correlate Sensor Results*, *Trust Values That Meet Expectations* and *Trust Values That Are More Accurate*.

#### 5.3.3.1 Own Research

Participants described researching devices when attempting to determine their trust. This research included reviewing online forums and public perception, *“it would be partly its public reputation, I would do some online research into how good it is”* (P4).

Participants also described researching online, with their research focused on more general forums due to lack of ME/CFS-related websites known to the participant, *“I use it on a general one because I had actually looked for ME/CFS sites looking at technology and I wasn’t able to find much information out there on those specific things which is a pity because that would be useful”* (P4). P5 described having used public forums to understand values that they did not feel they could trust,

*I know with the Garmin if I have a problem I’ll go on the forums and sometimes, like mine was measuring that I was doing over a thousand minutes of intense exercise a week which I don’t do at all. I was like that’s really strange and it’s based on your heart rate, so I was like is it measuring my heart rate wrong. So I went on the forums and some people had found that software updates had affected their measurements.* (P5)

#### 5.3.3.2 Compare and Correlate Sensor Results

Participants responses included in this theme relate to participants attempting to use the values presented by the sensing technology and correlating them with their own records, other devices or feelings of a particular recording session, *“I have things like I track my steps so I can literally walk around and count my steps and see if that matches.”* (P4). Participants described feeling their pulse and measuring it against a clock as a means of determining how trustworthy the values provided by a device might be, *“I’ve compared it against other recording devices at the same time but that’s something that is easy to do because it’s something tangible that you can feel your pulse*

*and you can measure it against a clock” (P1). Others compared values across devices, “I’ve got a blood pressure monitoring, and it does pulse and all the rest of it as well, and that seems to be, I wouldn’t say erratic but it doesn’t seem as steady as the Fitbit is showing up” (P7).*

Participants highlighted that this was not as easily done with other types of values, *“with anything cognitive it’s much more difficult, so I think that makes it more difficult because you have no gauge yourself, especially with Chronic Fatigue because how you feel, is often very different” (P9)*

#### **5.3.3.3 Inherent Trust in Technology**

This theme encompasses responses where participants have described trusting that the device is recording what it is supposed to measure due to some inherent trust that the device or technology is doing what it is meant to, *“I suppose I’m very trusting of what technology can do. I think that’s probably the reason. I don’t doubt its capability of doing things” (P7).*

Similar responses were seen where participant had no reason not to trust the technology, *“I’ve got no reason to distrust it because I don’t know... On the face of it I trusted it.” (P1).*

#### **5.3.3.4 Trust Values That Meet Expectations**

In this case participants responses are a kin to trusting values more than meet their expectations, which is also described in Section 5.3.2.3. However, this expectation of how data should behave appears to play a role in users’ trust in sensing technologies.

For example, when asked if they trusted the values that they had been shown, participants highlighted not trusting specific pieces of data because it did not match what they expected, *“I think the mood ones I didn’t trust as much because when I would rate my own mood sometimes it would be quite close to what I rated it and other times it would be way off” (P5).*

This way of building trust may be a result of confirmation biases, with users more likely to trust values that meet their prior expectation of how data should behave.

#### **5.3.3.5 Trust Values That Are More Accurate**

P3 and P8 made references to trusting devices more if they were more accurate. P3 gave a general statement about trust and accuracy, *“I’d probably trust it more if I thought*

*that everything I was given was accurate” (P3). P8 specifically mentions differences in accuracy between medical and non-medical devices, suggesting that users have more trust if something has been medically certified, and that medically certified means better accuracy, “You warned me that this was not medical, this was not medical grade. It was something still in development. So I guess the more accurate it was the more I would trust it” (P8). Therefore, devices which users perceive to be less accurate may result in a perception that the device and the data presented is less trustworthy.*

## 5.4 Discussion

### 5.4.1 Designs That Consider The Effects A Condition May Have on Users

Due to their condition participants in this study had specific issues coming to terms with the technologies and in being able to begin to analyse the data. For those with preexisting conditions who are seeking to make sense of their data so that they can better understand their symptoms we suggest that it is of the utmost importance that designers, developer and researchers include their intended user group in each stage of the research.

One limitation that we encountered in this regard was the availability of remote tools that supported collaboration with participants in designing interface, where there may be limited ability to travel or where participants’ regional distribution may make no one location suitable for in-person group research, such as participatory design (Muller, 2002). Although previous research has looked at the design and benefits of remote collaboration tools for participatory design (Heintz, Govaerts, Law, Holzer, & Gillet, 2014), currently available online tools appear more geared towards teams of developers or designers (e.g. Miro<sup>2</sup>). Although there are likely to be certain issues in this regards, such as varying technical abilities of participants and possible limitations of transferring something that is usually done in-person with items such as post-it notes. Being limited to only being able to carry out design research with participants who can manage to all be in one place at the same time detracts from the benefit of enabling them to participate in the design of tools made for them. Otherwise designers should spend significant time testing UI designs using alternative methods with the intended user group before making their tool widely available.

---

<sup>2</sup><https://miro.com/>

### 5.4.2 Supporting longer-term tracking

Participants in our study suggested that they were keen to track longer-term trends within their data, with a suggestion that periods of data up to 24 months would be of interest. At the time of writing, Fitbit does not provide a simple way for users to access this data. To review data a user must navigate through their dashboard day-by-day to find a days' data that is of interest (see Figure 5-6). A work around is to manually edit the URL to access the data for a specific data<sup>3</sup>. Users can download the previous 31 days' worth of data from the website, available in XLS or CSV format. If the user wants to access an extended range of data they must request all of their data. This exported data includes *all* account data, including things like: the times passwords were reset, times when the account email address was changed, friends list. On initiating a request for this data the user will be sent an email with a link to confirm they want to download the archive, clicking this link then begins the process of Fitbit pulling their data together and they can then download the entirety of the data when it is ready. Fitbit notes this could take anywhere from 30 minutes to several days depending on the amount of data. Fitbit provides the majority of this archive data in JSON format. Exporting extended data can be seen in Figure 5-7.

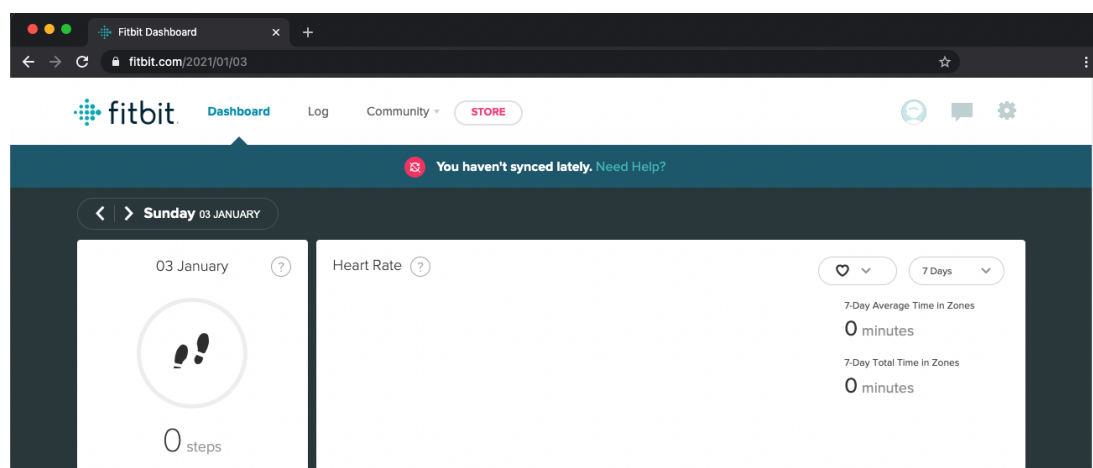


Figure 5-6: Fitbit dashboard showing one-day forward and back data navigation

Although it is good that people are provided with some access to their personal data, it limits the usefulness of the PI tool to limit users to accessing only subsets of data when they might be interested in much longer term trends. Additionally, one participant in this study who had download their data from Fitbit suggested that the format in which it is provided (JSON) is not the easiest to use. Thus a secondary issue related

<sup>3</sup><https://community.fitbit.com/t5/Other-Charge-Trackers/How-far-back-can-I-look-at-my-data-online/td-p/3456467>

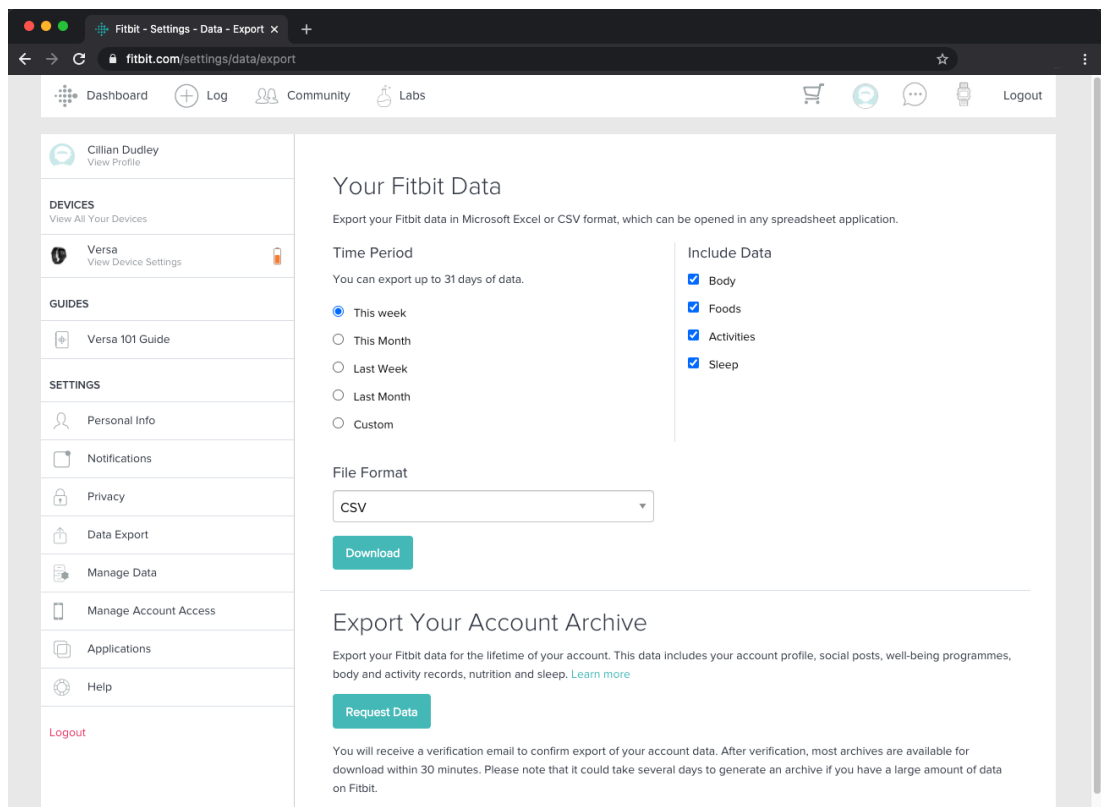


Figure 5-7: Fitbit user data export page - showing the two possible export options up to 31 days and account archive

to providing easier access to data for longer periods of time is that the data should be provided in a format which is easily manipulated and analysed by the users who are using the platform, not just technical or data-savvy users.

So although previous studies have highlighted that getting people to track over longer term may be difficult (Rapp & Cena, 2014), users also appear to be unable to actively engage with this longer-term data in a simple and intuitive way. Designers should consider both encouragement to track longer term but also design PI tools with longer-term tracking in mind and provide means to easily navigating these extended periods of time. If unable to provide the means to analyse data within a PI tool at least designers should enable users to access their data in a suitable format that they will be able to manipulate themselves. Those building tools to integrate data from multiple sources on behalf of users should provide mechanisms to do any necessary data formatting on the user's behalf, thus lowering the barrier to entry to these tools.

#### **5.4.3 Supporting target-setting and comparisons when previous data may not be available**

Participants struggled to know how to go about target setting or comparisons when they had no data to compare it to. This suggests a potential limit for any users who are attempting to use PI tools for diagnostic or deterministic reasons as unless they have been tracking data prior to being diagnosed with an illness they will be unable to determine what might be considered 'normal'. Additionally, participants highlighted the individual nature of symptoms encountered by people living with ME/CFS. Thus there is unlikely to be a one-size-fits-all approach; presenting information in relation to other users being unrealistic and potentially problematic when attempting to track, compare, or set targets in relation to someone who is not living with ME/CFS.

Further research is required to determine the best way of presenting and enabling this type of data for comparisons and providing means of setting healthy and realistic targets, if that is something that the user wishes to do. Feustel, Aggarwal, Lee, and Wilcox (2018) found that participants were able to use cohort data to determine averages for comparing themselves to a relative baseline. They also highlight that need for these baselines to be considered in the context of grounded research and "being mindful of the potential to focus on data that promotes the "status quo" for an individual" (Feustel et al., 2018). When considered within the context of individuals with chronic illnesses, a consideration of what a suitable cohort for comparison is requires further research such that the choice of data does not exacerbate problems for the user and that the cohort's data is appropriate to make comparisons with.

#### 5.4.4 Supporting a variety of data presentation formats, varying levels of data knowledge, and technical knowledge

Participants expressed a desire to be able to analyse their data on screens larger than a mobile device, with participants suggesting that tablet- and desktop-size screens were preferable. Raptis, Tselios, Kjeldskov, and Skov (2013) found the users who used a screen size larger than 4.3" performed better during information seeking tasks, in this study participants used screens varying from 4" to 5.8" ( $M=5.1, SD=0.52$ ). Highlighting that such information seeking benefits may not translate into groups with conditions that make information seeking difficult in the first place. Alternatively further work could be undertaken to investigate how to appropriately design visualisations for mobile displays, related to common time-frames that users are likely to be interested in analysing and providing appropriate visualisations to support these time periods (e.g intervals of 72 hours, 7 days, 6 months as seen in prior work (J. Meyer, Kazakova, Büsing, & Boll, 2016)). This is of particular importance when considered in relation to tracking longer-term trends in data, as the amount of data to be visualised increases in response to the larger time frame. An alternative or complimentary approach may be to consider alternative interaction modes, such as voice, to reduce the need to search using touch on mobile devices. Recent research has suggested that voice interactions can help in overcoming limitations related to smaller screen-sizes and manipulating data (Young-Ho Kim, Bongshin Lee, Arjun Srinivasan, & Eun Kyoung Choe, 2021). However, this may not immediately reduce the amount of data that is finally presented, which requires further consideration and research to understand what modalities might improve data analysis.

Participants' levels of knowledge in relation to data analysis varied, as such participants felt that some of the ways that were used to present data were too technical and potentially not relevant, which has been seen previously (Macleod et al., 2013). Ensuring that all users can benefit from recording and analysing their own data suggests a need to provide multiple ways of presenting information to users. Too simplistic presentations may be considered to provide no benefit to people who are keen to dive deep into their data. Too technical, detailed or elaborate data presentations could result in those less familiar with data analysis not being able to make a first effort to better understand themselves through their data.



#### **5.4.5 Providing information about the basis on which the data is formed**

Participants in this study showed both an awareness of, and interest in knowing, that the tools and the data provided has been grounded in prior research. Although this conflicts with our previous results which seemed to suggest that users were not concerned about the scientific basis of produced data, the suggestion remains consistent in both cases. We suggest that developers of novel sensing technologies make it clear what relevant research their device has been based on and make this information available in a manner that can be consumed by members of the general public. This will allow them to increase levels of confidence that there is some research behind the device and that the values that are produced by the device have specific levels of accuracy or reliability. Additionally, manufacturers may like to provide information related to benefits and potential pitfalls of using their technologies for the sake of transparency, ensuring that any potential ethical concerns are also addressed. These suggestions should also address issues related to implicit trust in technology by providing users with an open and honest explanation of the technology.

### **5.5 Limitations**

The participants in this study were self-selected and therefore may not represent the entire population of people living with ME/CFS, CF, or ICF. In particular symptoms of fatigue or malaise and the variation between people can be quite personal. The study does not account for the fact that people living with daily debilitating fatigue may not make use of the technologies as easily, or may not have the energy to be able to analyse the data. Only nine participants took part in this study. A larger group of participants could provide additional insights, and more generalisable results.

### **5.6 Conclusion**

In this chapter we presented the qualitative results of a three-week in-the-wild study involving nine participants who are living with ME/CFS, CF, or ICF. We have presented design considerations, challenges, and threads for future work, one of which (trust) we explore further in Chapter 7. We find, in line with previous research, that the way data is presented in applications can have negative impacts on users' conditions. Therefore, our design considerations suggest that it would be beneficial to provide users with tools designed to take into account the effect that their condition may have on their ability to record and interpret data. Of particular interest is our finding that des-

pite previous research suggesting that people needed further encouragement to track longer-term (and we do not suggest that this is not a challenge), participants in our study appeared to acknowledge the need for longer-term tracking and showed a willingness to do so. Therefore, we present a design suggestion that suggests those tracking chronic illnesses may be interested in longer-term tracking, which currently is not often supporting by tracking tools. We present a design challenge related to supporting the variation of technical and data knowledge. We also highlight a challenge that those who take up tracking to monitor symptoms of their condition may not have data from prior to diagnosis and suggest that further research is required to better understand how to facilitate them in making appropriate comparisons for their current state with a consideration about what an appropriate set of data for comparison might be. Finally, we provide a suggestion that developers of sensing technologies should provide their users with information about how the collected data is formed, and whether it is backed up by prior research.

## Chapter 6

# Encouraging Exploratory & Enquiring Approaches: Comparing the Insight Feed and Query Area Interfaces

## 6.1 Chapter Overview

In Chapter 3, Section 3.4.3, we described the challenge of encouraging exploratory and enquiring approaches to data analysis. This challenge centres around fostering engagement with, and exploration of, data that people have collected with the intention of finding answers and developing understanding of themselves (e.g. I have fewer symptoms of X when I do Y) rather than providing deterministic diagnoses (e.g. Do this and you will be cured). Rather than suggesting that a specific technology, or amount of data, can provide a solution or ‘cure’ to a problem, exploration and enquiry of recorded data may enable people to understand, for example, actions which result in exacerbating symptoms or situations that may lead to a relapse of a condition.

In this chapter we present results from the analysis of data related to the use of the Insight Feed and Query Area features in the Cognizance Tracker application, provided to participants during the study described in Chapter 5. Rather than providing simple aggregates and charts of recorded data which place a burden of analysis on the users, interfaces and mechanisms for providing people with alternative means of reviewing their data to improve sensemaking and tame complexity, have been studied previously. An example of such work by Epstein et al. suggests that providing users with ‘cuts’, a subset of data with some shared feature, of their data may provide those less experienced with self-tracking of a means of seeing high-levels patterns and reducing the burden of data synthesis (Epstein et al., 2014). Previous work has suggested that identifying significant connections between personal data may be beneficial in promoting self-understanding and behaviour change (Bentley et al., 2013). In this study we aim to understand whether the Insight Feed and Query Area features enable users to explore their data in additional ways that are beneficial, such as the ability to more easily make sense of their recorded data, bringing awareness to information that they would not uncover as easily using more traditional means of analysis and the challenges they may face.

In designing interface mechanisms to support engagement with and exploration of data generated by the Cognizance Tracker application, we drew inspiration for the Insight Feed feature from the many modern media content platforms which attempt to deal with information overload using algorithmically curated ‘News Feeds’ (e.g. Facebook, Twitter, Reddit). These news feeds attempt to present users with information which is mostly likely to be of interest and filtering out those which it considers to be of least interest (Bernstein, Hong, & Kairam, 2010; Koroleva, Krasnova, & Günther, 2010; Mills, 2011). This has been combined with features from commercially available

tools for analysing multifaceted data, such as Exist.io<sup>1</sup>, which has been previously used in PI research (see S. L. Jones and Kelly (2018)). The Exist application enables users to collate data from disparate Personal Informatics data sources such as Google Fit, Apple Health, Fitbit, Twitter, Spotify, and others. Exist then finds correlations within a user’s data and provides insights such as “Monday is your most active data of the week” and “You have a better day when it’s the weekend”. The correlations are graphed and presented to users with a star rating reflecting the confidence (p-value) of the correlation and a percentage value reflecting the strength of the correlation (see Figure 6-1). The use of Exist has been explored by S. L. Jones and Kelly (2018) to understand the data sensemaking challenges faced in using such a system.



Figure 6-1: Exist’s Correlation Feature

The Insight Feed for Cognizance Tracker was developed as a way of bringing attention to data without participants having to manually search for insights themselves. The Query Area on the other hand provided users with a way of analysing recorded data by selecting from a choice of predefined analysis types, analysis variables and a date range. For example, users could select a query to show them ‘*CORRELATIONS* between *FATIGUE RATING* and *HEART RATE* from *START DATE* to *END DATE*’ or select a query to show them the ‘*TREND* of *HEART RATE* from *START DATE* to *END DATE*’. The types of insights generated and query types available are based around a subset of the types of visualisations that members of the quantified-self movement have previously been noted to use to make sense of their data (Choe, Lee, & Schraefel, 2015): correlations, extremes, outliers, trends, and variability. Trends and correlations were also suggested as analysis types by participants in the study presented in Chapter 4

<sup>1</sup><https://exist.io/>

(see Section 4.3.3.1).

The features discussed in this chapter aim to answer the following research questions:

**RQ1** Does an ‘Insight Feed’ of pre-computed insights, encourage further exploration and sense-making of data?

**RQ2** Does providing a predefined set of possible queries, as in the Query Area, enable people to better explore and make sense of their recorded data?

To answer these questions we recorded and analysed participant application usage data. Application usage recorded includes interactions with the Insight Feed, Query Area and the general Overview of recorded sessions. We also analysed participants’ responses that related to these specific features during the post-study interview questions, asked as part of the interviews described in the previous chapter.

## 6.2 Insight Feed

The Insight Feed (see Figure 6-2) displayed automatically generated statistical insights to participants in the form of ‘cards’, examples of these cards can be see in Figure 6-3. The cards appeared in a scrollable news feed-style interface. Each card presented users with information about one of the following insight types: Correlations, Extremes, Variability, and Trends, found in their collected data. Each card was composed of a title which display the variable(s) being measured, a text description of the determined insight, a thumbs up and down icon, an info icon, and a ‘More’ button. In the case of Correlation and Trend cards, a star rating was used to present confidence information. The thumbs up and down icons could be used by the user to rate the insights that were presented to them and to highlight insights that they found useful or not. Following this participants were also optionally asked to rate the insights on a scale of 1 to 5 for the following dimensions: how interesting they found the insight, how much they trusted the insight, how useful they found the insight, how well they understood the insight (see Figure 6-4a). The information icon presented descriptions of the variables on that insight card (see Figure 6-4b).

Insights were generated daily, on a web server, and were later downloaded by the application on participants’ phones. Participants received a notification on their phone telling them when new insights had been synchronised to the application. Insights were generate such that participants received an insight if it differed from the previously generated insight for the same metric(s) or at least 7 days had elapsed since that specific insight had last been generated.

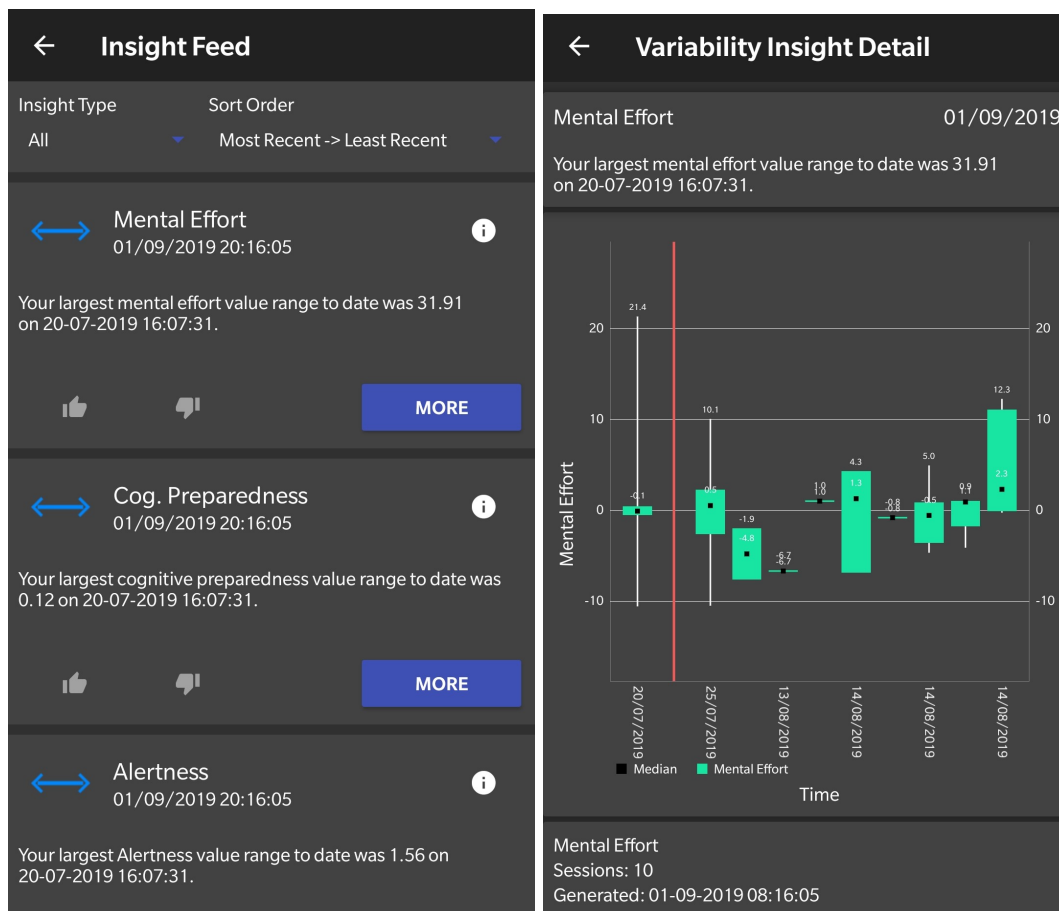
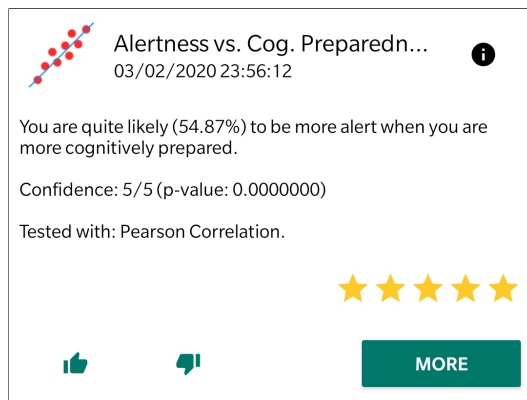
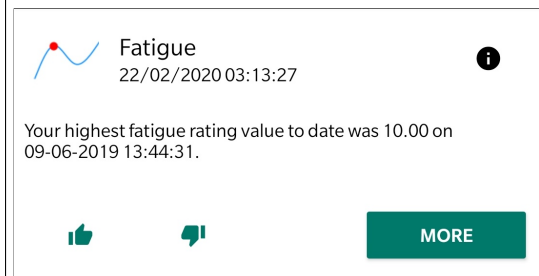


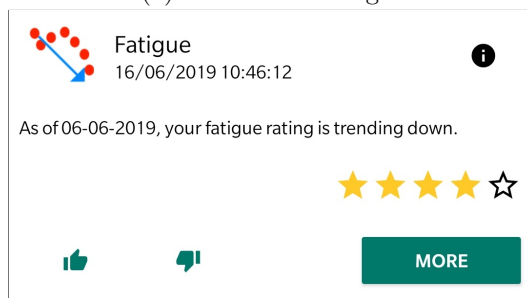
Figure 6-2: The initial Insight Feed page (left) and the Insight Detail page of a specific insight accessed via the ‘More’ button (right).



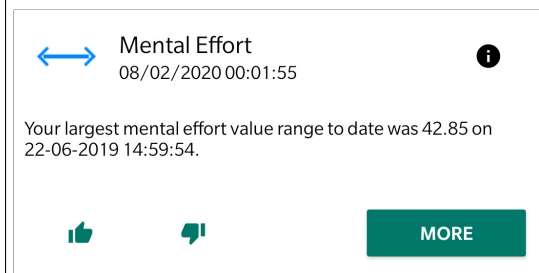
(a) Correlation Insight



(b) Extreme Insight



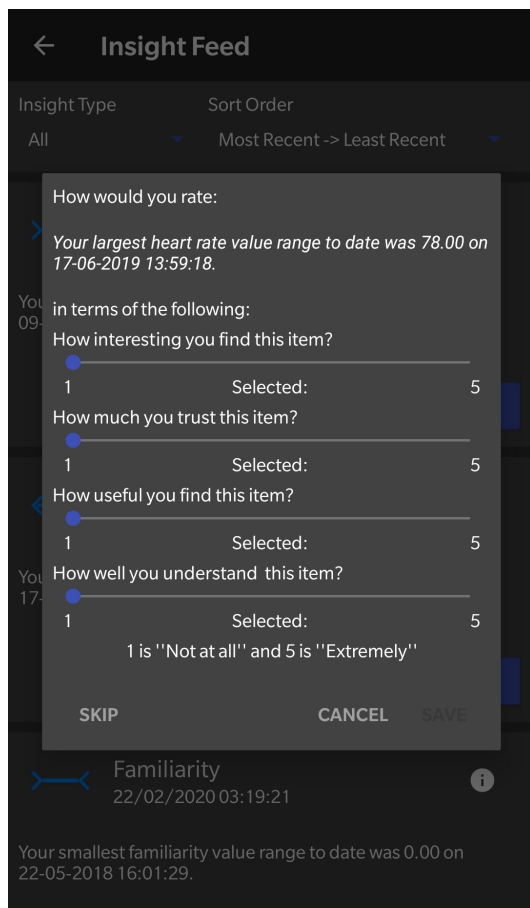
(c) Trend Insight



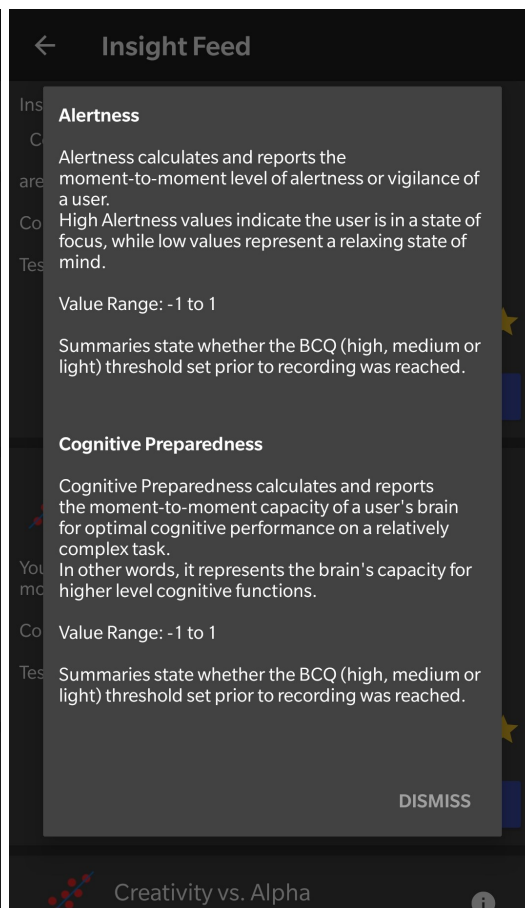
(d) Variability Insight

Figure 6-3: Example Insight Cards





(a) User insight rating dialog



(b) Insight variable information

Figure 6-4: Insight rating dialog shown to users (left) and Insight variable information shown to users (right)

### 6.2.1 Correlations

In Chapter 4, some participants described being able to find correlations in the data whilst others did not. We therefore suggested that it may be beneficial to raise these correlations to users’ attention and attempted to do so with these features.

P#	Total Correlation Insights	Corr. Insights ( $p \leq 0.2$ )
P1	1140	865
P2	724	529
P3	971	709
P4	1140	880
P5	570	456
P7	842	603
P8	989	723
P9	1127	814
P10	608	433

Table 6.1: Total and Sent Correlation Insights by Participant

Correlations were generated against all metrics resulting in 105 possible correlation combinations. Correlations deemed to have a p-value of less than or equal to 0.2 were synchronised to a participant’s app such that they appeared in the Insight Feed. A p-value less than or equals to 0.2 is generally not accepted as a good level of statistical significance, however, this has the potential to show tentative correlations. Precedent for this level has been seen in Exist.io (S. L. Jones & Kelly, 2018). The total number of generated insights versus the number of insights sent to each participant, on the basis of  $p \leq 0.2$ , can be seen in Table 6.1.

r value	Likelihood Phrase
= 1.0	almost certain
$\geq 0.7$	very likely
$\geq 0.5$	quite likely
$\geq 0.3$	a little bit likely
$< 0.3$	not likely

Table 6.2: Caption

The text of a Correlation insight was composed by combining “You are” with a likelihood phrase based on the measured correlation coefficient (see Table 6.2), the phrase ‘to be more’ or ‘to be less’ depending on whether the correlation was positive or negative, and a more fluent expression of the metrics being compared. Previous work has highlighted that presenting information in natural language has benefits of increasing awareness self-influencing factors, promoting the creation of behaviour change

strategies and increasing understanding of complex relationships (Bentley et al., 2013). Additionally, this style of presentation is already in use in commercially available systems such as Exist.io (S. L. Jones & Kelly, 2018). Examples of generated correlation phrases include:

‘You are a little bit likely (36.02%) to be more creative when you are experiencing more pleasurable emotions (valence).’

Creativity (EEG) & Valence (Affect Grid), r: 0.3602

‘You are quite likely (80.31%) to be more creative when your alpha values are higher.’

Creativity (EEG) & Alpha (EEG), r: 0.8031

‘You are a little bit likely (54.57%) to be more cognitively prepared when you are more meditative.’

Cognitive Preparedness (EEG) & Meditation (EEG), r: 0.5457

### 6.2.2 Trends

In addition to correlations, participants from the study described in Chapter 4 expressed a desire to use trends to determine how values were progressing over periods of time.

P#	Total Trend Insights	Trend Insights ( $p \leq 0.2$ )
P1	140	111
P2	97	79
P3	125	93
P4	140	114
P5	69	56
P7	100	76
P8	124	105
P9	139	118
P10	80	59

Table 6.3: Total and Sent Trend Insights by Participant

Trends used all of the data to date to determine if the values of particular metric were trending up or down by calculating a trend line for each of the metrics and then determining the p-value of this trend. Only trends with a p-value less than 0.2 were sent to participants. The total number of Trend insights generated and made available to participants can be seen in Table 6.3.

‘As of 26-06-2018, your Meditation is trending up.’

Meditation, Trend Insight

### 6.2.3 Extremes

P#	Total Extreme Insights	Min Insights	Max Insights
P1	126	65	61
P2	119	58	61
P3	127	62	65
P4	118	55	63
P5	131	65	66
P7	143	67	76
P8	122	59	63
P9	126	62	64
P10	126	60	66

Table 6.4: Total Extremes Insights by Participant

Minimum and maximum values were extracted from participants’ data and used to generate Extreme insights for all metrics. After the first recording participants had a Minimum and Maximum insight for each metric. After this only minimums and maximums that were higher or lower than the previous insight’s value were sent to participants’ Insight Feed. This type of insight presented an indication of when a lowest or highest value to date had been recorded. Examples of extreme insights are as follows:

‘Your lowest Creativity value to date was -0.69 on 02-06-2018 16:51:10’

Creativity, Min Insight

‘Your highest Meditation value to date was 100.00 on 07-02-2018 12:20:21’

Meditation, Max Insight

### 6.2.4 Variability

Variability insights were generated by calculating the difference between the lowest and highest values per session. In reality, participants were presented with the range of values rather than the statistical variability of their data. After the first recording session participants had a Variability insight for each metric, after this only metrics with higher or lower ranges were sent to participants.

P#	Total Variability Insights	Smallest Range	Largest Range
P1	100	43	57
P2	111	61	50
P3	110	57	53
P4	91	46	45
P5	105	52	53
P7	102	42	60
P8	101	47	54
P9	98	44	54
P10	97	44	53

Table 6.5: Total Variability Insights by Participant

‘Your largest Alpha value range to date was 17.87 on 26-06-2019 14:59:54.’

Alpha, Largest Variability Insight

‘Your smallest Creativity value range to date was 0.00 on 01-06-2018 17:24:56.’

Creativity, Smallest Variability Insight

### 6.3 Query Area

The Query Area (see Figure 6-5) provided participants with a way to conduct the following analysis types: correlations, extremes, trends, variability, and outliers, on their recorded data. Participants were able to see a graph presenting the results of the analysis. Additional relevant data was shown alongside the graph such as p-values, correlation strength, and the number of recording sessions included in the analysis. However, this section did not provide any further interpretation of the data. It was then left up to participants to interpret and analyse the data further, if they wanted to, such as to improve self-understanding, monitoring health and well-being.

### 6.4 Methodology

The analysis presented in this chapter uses data collected as part of the 21-day in-the-wild study described in Chapter 5, with 9 (M: 4, F: 5) participants whose ages ranged from 23 to 75 ( $M=45.67$ ,  $SD=17.69$ ). P6 is intentionally missing from the presented results, having withdrawn from the study. The study methodology received ethical approval from the university’s Research Ethics Approval Committee for Health (REACH) (Reference: EP 18/19 028). Data pertaining to this study is stored in accordance with the data management plan in Appendix F. Participants used an application which could

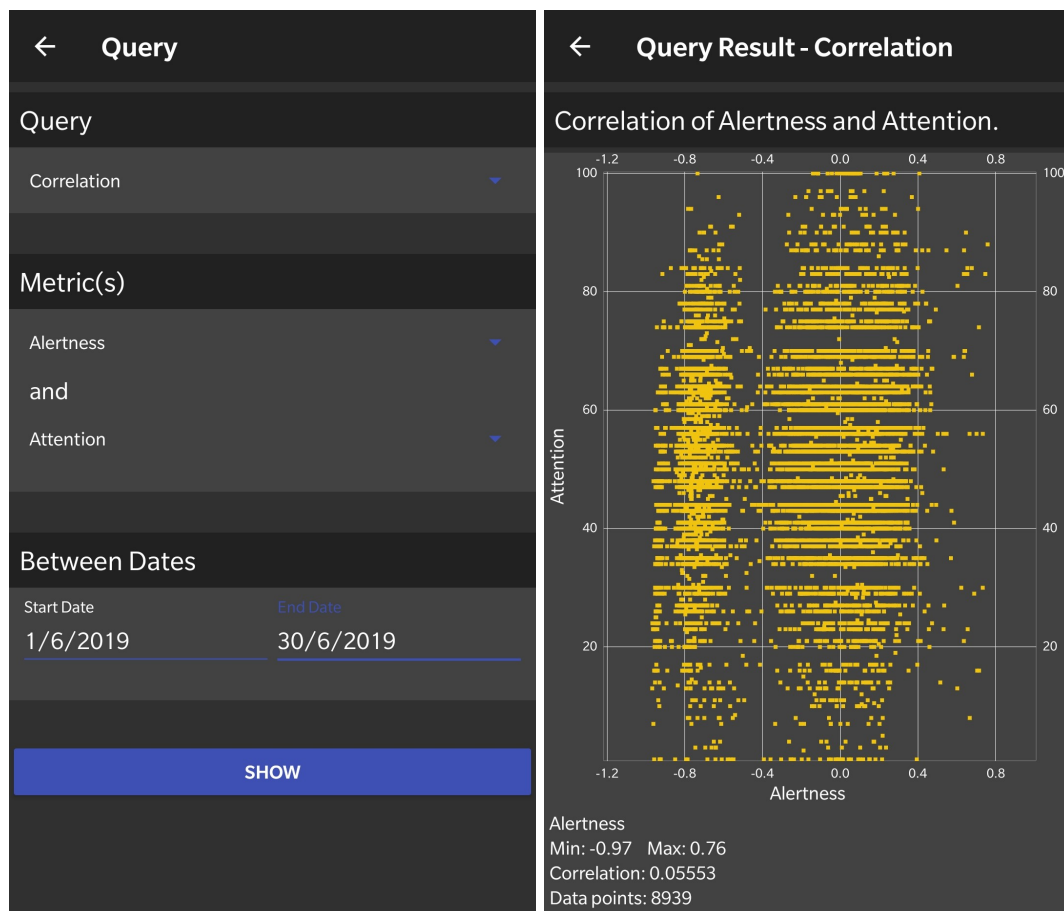


Figure 6-5: The left image shows the initial Query Area query selection screen. The right image shows the result of a Correlation Query result screen.

consolidate data from an EEG-headset as specific metric values, manual mood and fatigue ratings, and heart rate data captured via a Fitbit smartwatch. Additionally, the application recorded participants' interactions with the features of the application for post-hoc analysis. As part of the post-study interview conducted during the study participants were asked specifically about their usage of the Query Area and Insight Feed. The purpose of these questions was to understand how both features were used, why they were used and any difficulties participants had using either feature. If participants had made use of either feature, they were asked to describe how they used that feature.

For those who used the Insight Feed, they were also asked which type of insights they found most and least: interesting, trustworthy, understandable and useful. Participants were also asked when they were most likely to use the Insight Feed Area. For participants who made use of the Query Area feature, they were asked which types

Participant #	Age	Gender	Diagnosis
P1	44	M	ME/CFS
P2	75	F	ME/CFS
P3	61	F	ME/CFS
P4	36	F	ICF
P5	34	F	ME/CFS
P7	57	M	CFS
P8	56	M	CF (Cancer)
P9	25	F	ME/CFS
P10	23	M	CF (Ehler-Danlos Syndrome)

Table 6.6: Study 3 - Participant Demographics

of queries they used the most. They were asked if they were able to find specific data or insights that they were looking for, as well as how they used the data when they did find it. They were asked if there were any queries they wanted to be able to perform but could not, and if so, what they did instead. Participants were also asked in what situations/contexts they used the Query Area. If a participant had not made use of either feature they were asked why not and if there was anything that would have encouraged them to make more use of the feature. The purpose of these questions was to understand the reasons participants chose not to make use of these features and to draw out information that would enable us to understand what future designs of these features might benefit from incorporating.

Participants' interviews were transcribed in full and responses related to the features presented in this chapter were analysed in isolation from responses presented in the previous chapter. The relevant isolated portions of the interview transcripts were coded in full, using NVivo. Codes were generated using an inductive, reflexive approach. The codes were reviewed and initial themes were generated from codes with shared meaning, and the themes were then further refined. The results of this qualitative analysis are presented in the following Results section (Section 6.5), and further discussed in relation to four design themes in the Discussion section (Section 6.6) of this chapter.

## 6.5 Results

### 6.5.1 Feature Usage Comparison

To provide some initial context and as a basis for comparison, Table 6.7 shows the number of times that participants interacted with the Session Overview (which allowed participants to review their recorded data on a session by session basis using line graphs

P#	Interaction Count		
	Overview	Insight	Query
P1	5	<b>13</b>	7
P2	<b>7</b>	1	2
P3	<b>23</b>	19	3
P4	<b>14</b>	10	1
P5	<b>10</b>	7	3
P7	2	1	1
P8	<b>43</b>	15	4
P9	7	<b>11</b>	10
P10	5	<b>7</b>	2
Total	116	84	33

Table 6.7: Feature Usage - Interaction Count

and bar charts, see Figure 5-4), Insight Feed and Query Area sections of the application. For the purpose of our analysis an interaction is counted as any interaction related to a specific area of the app, within a ten minute window. Generally the Overview section was the most used area of the application with the exception of P1, P9 and P10 who viewed the Insight Feed most. The Query Feed was used the least by all participants.

P#	Mean Time Spent (mins)		
	Overview	Insight	Query
P1	1.047923	<b>1.544119</b>	0.409648
P2	0.303912	<b>6.642567</b>	0.114267
P3	<b>6.272659</b>	3.976718	1.489144
P4	<b>3.508177</b>	1.427033	1.060150
P5	3.158357	1.593519	<b>8.207133</b>
P7	3.700975	1.336267	<b>9.039450</b>
P8	2.254619	<b>3.993212</b>	1.651396
P9	2.014331	1.644377	<b>3.453448</b>
P10	2.695040	<b>3.604183</b>	2.072983

Table 6.8: Feature Usage - Interaction Length

Although not necessarily a direct indication of engagement (see Section 6.6.1), analysis of the average time spent during interactions with the three areas is shown in Table 6.8. Based on average time spent interacting with the areas, Insight Feed, followed by Query Area and finally Overview is the order by which participants spent most time engaging with these features. Four participants on average spent more time engaging with the Insight Feed area, followed by three participants engaging with the Query Area and two participants engaging with the Overview for the longest amount of time.



In the following sections we provide further analysis of the Insight Feed and Query Area, providing a quantitative analysis of participants interactions as well as drawing on qualitative data from interview responses in order to understand the benefits and challenges presented to people when using the Query Area and Insight Feed.

### 6.5.2 Using the Insight Feed

Participants' responses suggest two different ways in which the Insight Feed was used. P1 mentioned that they felt using the Insight Feed was beneficial in reducing the burden of data analysis:

*[R]ather than just leaving it just down to the user because there are so many parameters, rather than just leaving it down to the user to say, "What about this versus this", "I wonder what that's been doing?", having that sort of pre-made list to sort of have a flick through was quite useful I think. (P1)*

This was an intended use of the Insight Feed in allowing the user to gain an understanding of their data without having to perform the exploratory analysis themselves. However, another participant mentioned that they used the data obtained from reviewing the Insight Feed to inform further analysis and exploration, by reviewing the overall session to which the highest and lowest values (Extreme Insight) were from:

*I looked at the Insights and as I got more Insights and more information when I finished on Monday I went back through the Insights and looked at where was the highest, where was the lowest of things, what was that linked to? And then went back and looked at the review of the entire session to see what else was going on. (P3)*

This suggests an alternative way of drawing attention to points that may be of interest to the participants, but rather than being the endpoint of the analysis, this can be used as the starting point for further exploration. As one of the aims of this study was to determine whether the Insight Feed encouraged further exploration and enquiry in recorded data, we see that participants engaged further with session data after being nudged to look at the data based on specific pieces of 'interesting' information that were automatically generated.

Figure 6-7 provides further details by comparing participant interaction with different features of the Cognizance Tracker app by day of the study. Figure 6-8 shows interaction solely with the Insight Feed. Figure 6-9 shows similar detail but with interactions with the Insight Feed removed. In all of these figures interactions are given equal sizing

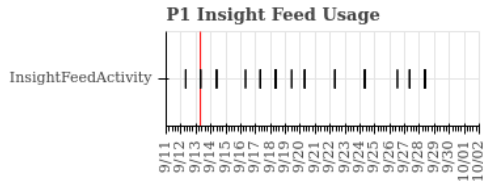
(regardless of duration of interaction) across participants but ordering of the events by time is maintained.

In Figure 6-9 we see that for the majority of participants' interactions tended to be around looking at the same type of insights. Or in cases where there are multiple insight types viewed, there tends to be a weighting toward one or two specific insight types. Also noticeable is that participants appear to view the Insight Feed without reviewing Insight details. There are noticeable gaps when the the Insight Feed data is removed. Which highlights that the Insight Feed itself was being used without participants further exploring their data using the 'More' button. This suggests that participants were still interested in reviewing the 'card view' of the data but did not always dig deeper into the information.

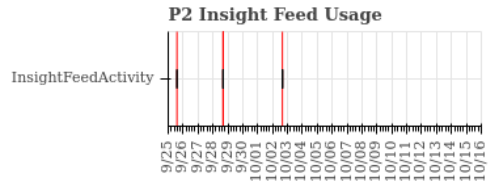
Figure 6-6 shows whether participants' interactions with the Insight Feed were via a notification or via the Insight Feed option from the application's main menu. The black lines show the interaction and the red vertical lines show whether that interaction was based on using the menu to access the Insight Feed. It would appear that, with the exception of P1, most participants in most instances, interacted with the Insight Feed by navigating to it by themselves from the menu screen of the application. This data is backed up by the qualitative data which suggests that P1 found the notification to be a significant factor in their use of the Insight Feed, *"having that come through was a good nudge to then have a sort of a scroll through"* (P1). Other participants did not make any particular comment about the notifications, which does not necessarily mean they were not useful but the quantitative data suggests that participants still mostly accessed the Insight Feed of their own accord. Previous work has highlighted the beneficial use of notifications as reminders or suggestions about tracked data (Epstein et al., 2017) but they have also been seen as a negative when extending beyond the period a user has been tracking data (Epstein et al., 2016). In this instance it does not appear quantitatively that the notifications had a direct impact on encouraging participants to explore their data in the moment, although the qualitative responses suggest they may provide a gentle nudge to explore data.

### 6.5.3 Utility of Different Insight Types

Participants had varied reactions to the types of Insights that were available. Generally, participants regarded the inclusion of Correlation insights as valuable. P1 described being *"suckered in by the correlation bit"* (P1) after receiving the daily notification for new Insights. P8 described finding *"correlations most understandable"* (P8). P3 however, felt that although the correlations were interesting to see, they were cautious



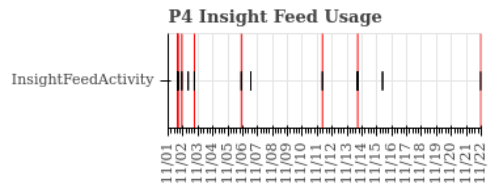
(a) Participant 1



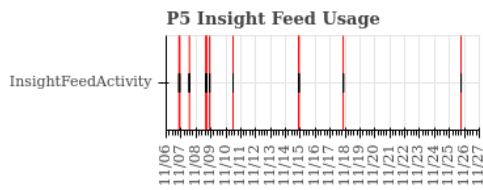
(b) Participant 2



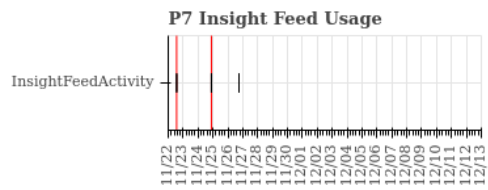
(c) Participant 3



(d) Participant 4



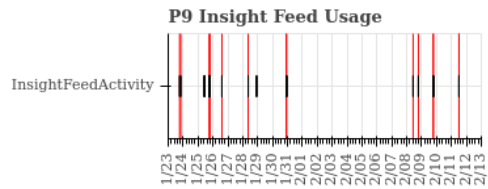
(e) Participant 5



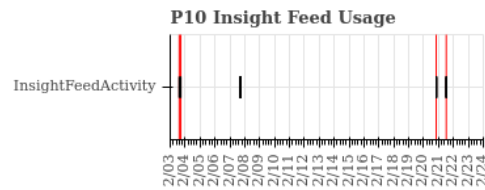
(f) Participant 7



(g) Participant 8



(h) Participant 9



(i) Participant 10

Figure 6-6: Insight Feed Initiation and Usage Over Time

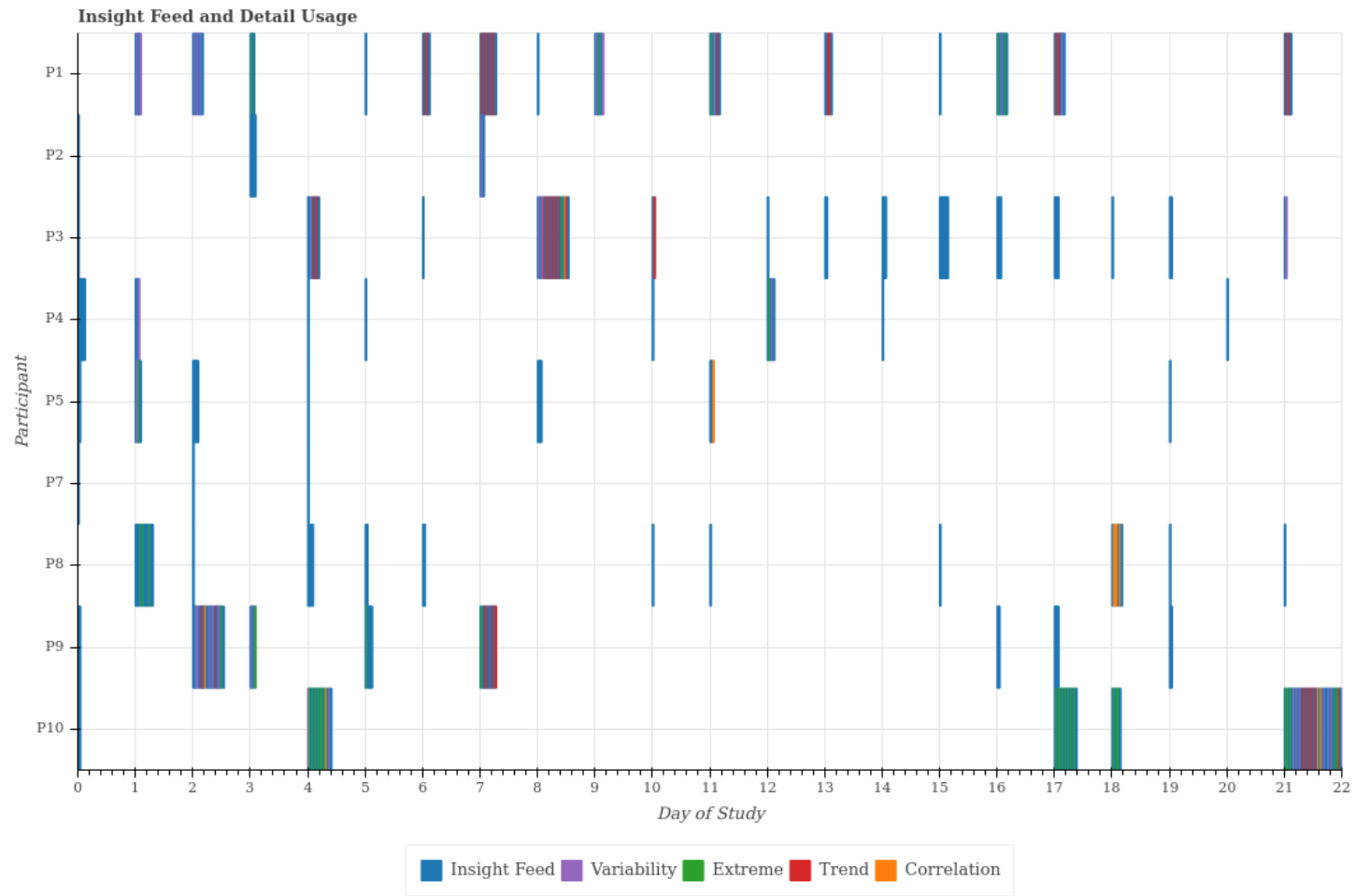


Figure 6-7: Insight Feed &amp; Detail Usage By Participant

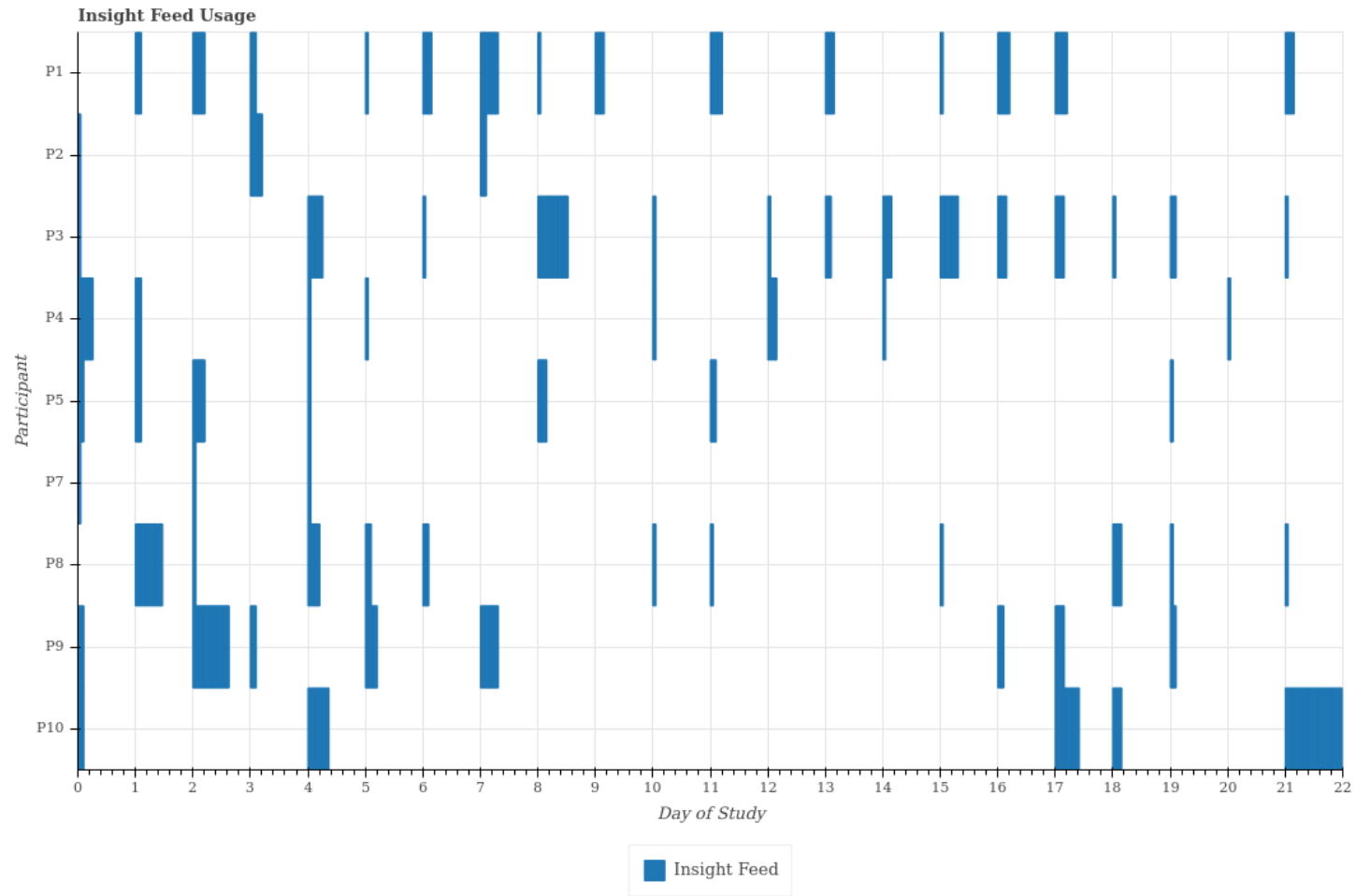


Figure 6-8: Insight Feed Usage By Participant

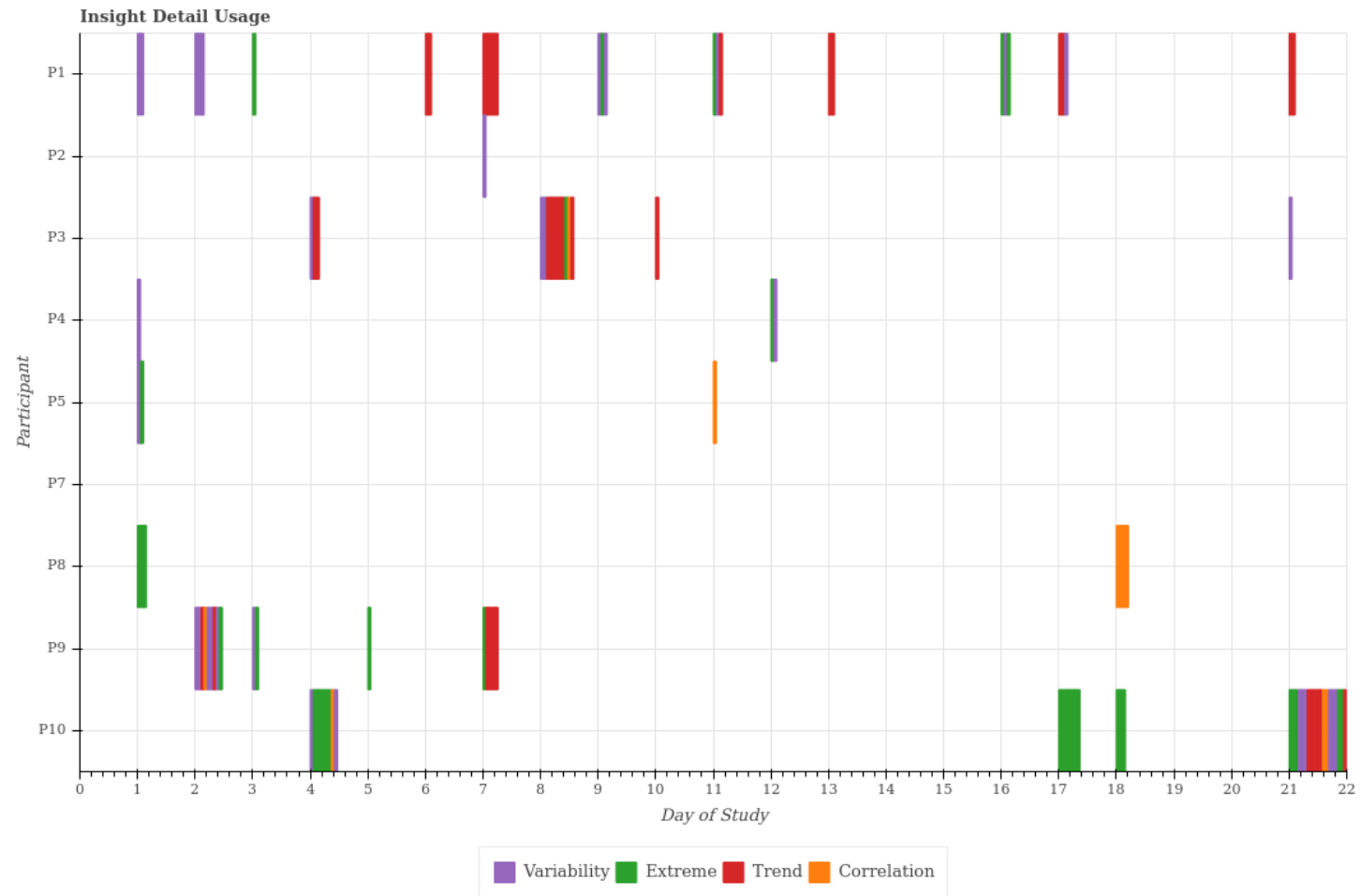


Figure 6-9: Insight Detail Usage By Participant

of how they should interpret what they mean:

*Correlations were quite interesting but I wouldn't think, "Ah that correlates, therefore..." but it was quite interesting to see what does. (P3)*

P1 felt that the Extremes insights were useful but only for metrics they were particularly interested in, rather than all Extremes Insights generally. They suggested that for *"any reports on the metrics I was most interested in, so the mental effort and the alertness, so if it was saying that I had highest value on a particular date then I would click and have a look"* (P1). Another participant felt that the Extremes insights were interesting but only with more data, *"The most interesting were things like where were your highest and lowest Alertness, Mental Effort... Extremes again I think I would want more data."* (P3). Other participants perceived the Extreme insights to be least useful and least understandable.

Comments regarding Trend insights were mixed, with one participant finding it useful, one participant suggesting it was the least interesting and another participant suggesting that it may be interesting but only with more data being acquired over longer periods of time. P5 highlighted a particular use case where they believed this would be helpful, saying:

*[O]ver time trends might have been interesting. I know with the heart rate sometimes I found it interesting, like, if I've switched a medication and I've been on it a couple of months then it's interesting to see if there's a trend that, kind of, correlates with that. (P5)*

However, one participant questioned whether the detected trends were meaningful or just as a result of normal variation. Another suggested that, because the recordings were for various different types of activities, the trends could not be easily related back to the specific activity causing the change in trend.

*I can't see what I can do with that when I've done lots of different activities at different times of the day and I didn't have an abnormally fast heart rate to start with. (P3)*

Variability insights were not commented on much during the interviews. One participant stated that they found them least useful due to not understanding what Variability was showing them and P5 suggested it was because their analysis was focused elsewhere and mostly made use of finding correlations:

*Variability I don't know if I would have found that very useful. I mainly*

*focused on, kind of, just looking at the individual graphs for each thing overall and then looking for correlations in other things, I didn't focus on the others much. (P5)*

The variety of responses and variability in which insights were considered helpful, useful or understandable suggests that the types of insights that users might find most utility in is marked by individual differences. This appears to differ depending on a person's ability to understand the type of insight, how willing they are to believe that the insight is being correctly measured, and their level of confidence in a specific type of analysis.

Table 6.9 highlights which insight types participants engaged with most, by navigating to the detailed view for that insight. Correlation-related insights seem to be of least interest to participants with the total number of correlation views accounting for only 8.8% of the total views and having 68.75% fewer views than the next most viewed insight type, Variability. This would appear to contradict prior work that highlights the eagerness of users to be able to detect correlations in data (Rapp et al., 2018) and our prior results in Chapter 4, Section 4.3.3.1, that also suggested that correlations were of interest to users.

Popularity based on totals does not quite tell the entire story though, with individual participants' most viewed insights varying. P10 showed a clear preference for Extremes, viewing 21 of them. Trends were more popular for P1, P3 and P9. P7 notably did not review any insights in more detail and based on the number of views engaged with the Insight Feed itself the least. As participants could see insight details from the Insight Feed (on each Insight Card) this analysis does not provide a full picture. Participants could potentially have read mostly Correlation insights from the Insight Feed and just not have viewed more details about them.

#### **6.5.4 Challenges with the Insight Feed**

A significant issue that participants encountered was the sheer number of insights that were generated. The number of insights generated for each participant can be seen in Section 6.2. Participants said that there were “too many permutations”, “too many breakdowns”, it was “too complicated” and there was “too much information”. P1 felt that there was *“an awful lot of permutations of how you can look at the data that's been recorded. Which is a good thing if you know what you're looking at and I think maybe that's where I struggled a little bit with some of it”* (P1). One participant, due to the number of generated insights, was not aware of all of the different types of insights being generated as they only saw certain ones in their Feed and other insights



P#	Number of Views				Total
	Correlations	Extremes	Trends	Variability	
P1	-	5	14	10	29
P2	-	-	-	1	1
P3	1	1	10	4	16
P4	-	1	-	2	3
P5	1	1	-	1	3
P7	-	-	-	-	-
P8	4	3	-	-	7
P9	1	4	6	6	17
P10	3	21	6	8	38
Total	10	36	36	32	114

Table 6.9: Insight Views By Insight Type and Participant

were not noticed when reviewing the Feed, *“I didn’t notice the filter function at the top. So when I was scrolling through and only seeing one or two types of insights I didn’t realise there would be more if I kept scrolling”* (P5). These comments suggest that providing a multitude of insights to review may have a negative consequence of overwhelming those who do not initially have an idea of what insights they are able to discover. Additionally, a high volume of insights being generated may result in some insights being lost in the noise of other insights.

Participants who engaged with the graphs, showing more detailed information about a particular insight, expressed that some of the charts were difficult to interpret. One particular issue that arose was with bar charts and correlations between continuous and non-continuous data. The difficulty in interpreting the graphed data caused participants to ignore those Insights completely.

*The way the data is presented there I didn’t find so easy to make sense of because it’s just giving you four bars of lines. So I tended to just scroll past those because I was like, “I’m not sure if I’m going to get much useful out of that graph” but then other graphs that would be presented in different ways I would engage with those a bit more.* (P5)

P9 expressed a preference for the text-based Insights over reviewing graphs. Referring to the general Insight Feed and cards’ text descriptions P9 said:

*[I]t’s then when you click on the more bit that I then got a bit more confused cause I had the graphs again. I think heart rate was okay, because it’s a number. Whereas anything that sort of went up and down, or, you know*

*when you got those fuzzy bars that go in between values, it kind of was less easy. So, I think it was kind of the layout of the graph that I struggled with, rather than the specific value, like the dataset. (P9)*

There is perhaps a balance to be achieved in regards to allowing users to further explore the data that the insight is based on in the form of a graph and encouraging users to go and explore the original data. Participants were already able to see the insight described in text and perhaps this would have been sufficient, rather than providing additional information in the form of graphed explanations of how that insight looks.

However, this solution is somewhat overshadowed by participants' feelings about the insights themselves. Participants felt the Insights were understandable but they questioned the usefulness of the Insights provided and were not able to come to a conclusion after reviewing the insights. P1 suggested that the Insights *"all presented the information in the way that was useful, as understandable, it was just don't know whether I found it useful or not"* (P1). Another participant was glad they *"could look at it but the end result hasn't been that useful because it's not been particularly conclusive"* (P9). This again suggests that an Insight Feed-style method of bringing attention to statistically significant data, doesn't provide solutions or answer questions. This style of exploration may only lead to more questions without providing answers to users' questions, such as "Why?" or "What next?".

Some participants highlighted that they were surprised that certain Insights did not show up how they expected. For example P3 said, *"I was a bit mystified about why a large Mental Effort didn't appear to go with a large amount of Attention and Alertness and Cognitive Preparedness"* and also found conflicting sounding Insights appeared which caused them to question their understanding of the results:

*Something else said something along the lines of more likely to be alert when I was fatigued, and I thought does that mean... That's a bit odd because you think if you're fatigued you wouldn't be alert but does that mean I'm trying much harder to be alert when I'm fatigued and therefore it takes more of an effort. (P3)*

Participants who felt that the results were not as expected noted that this may be due to the way the data was collated, *"the correlation didn't really show what I expected it to just because of the way that the data was collated"* (P1). The current iteration of the application generated the Insights based on all of a participant's available data. This was brought up by another participant suggesting that as a feature it would be helpful to be able to filter insights based on time and/or activity, or by providing the

ability to run Insight analysis on specific subsets of data.

One participant misinterpreted the meaning of the statistical star rating system of confidence. The participant believed that it appeared to be more like gamification and encouraging users to meet certain targets, which they were not keen on as a feature of tracking data, *“if you go further down your emotional response, you got two stars out of five”* (P10). Although this feature is directly copied from the Exist.io application, it highlights the need to clearly highlight features and consider ways which users might misinterpret symbols.

### 6.5.5 Using the Query Area

Participants used the Query Area to compare data across days, to explore data further as a result of seeing an Insight in the Insight Feed, and after recording sessions to see changes in their data. Using the Query Area as a means of further exploring Insights was mentioned by P3 and P5. P3’s description suggests using the Query Area to dive deeper into the data highlighted by the Insight Feed:

*[I]f I got something on the Insight Feed that said, “Your highest or lowest measurement was on this date.” I’d use the Query to go back and say, “OK where is the data from that date?”* (P3)

P5, however, used their assessment of how ‘interesting’ they found an Insight to be to guide their exploration of other metrics or data points:

*[I]f something had come up in the Insight Feed and I looked at the graph and thought, “Oh that’s maybe interesting” then I would then go and do a query to compare it with something else.*

P9 suggested that the Query Area was better suited to obtaining *“a quick answer because the Insight Feed, you kind of have to go through a bit more... Query I found a bit more direct”* (P9). Participants described using the Query Area as a way of gaining an understanding that the data behaved in an expected way, *“see that it’s recorded and see that it made sense to me that I was working”* (P7). Another participant suggested that a useful use case for the Query Area was to explore data facets that were of interest to them, *“it allowed me to go off and explore the data that I was interested in”* (P1). These responses suggest that participants used the Query Area when attempting to access specific information of interest them, or to gain a deeper understanding of existing questions that they wanted answers to.

Conversely, when participants who did not use the Query Area were asked about this,

one participant stated that they “*didn’t have a question I wanted to ask it*” (P8). This suggests that the Query Area may not be useful to those who do not have some initial question, idea or hypothesis about what their data may reveal. Participants did not tend to generate queries unless they had predetermined that the results may be insightful or meaningful to them. However, based on the participants’ responses described previously, it may be possible to elicit some curiosity to explore the data using alternative means. One such alternative could be prompts from insights generated by the Insight Feed leading to further data exploration (see Section 6.5.2). Another participant reported that they did not make use of certain queries because they did not find the results useful when browsing the Insight Feed, so did not feel that they would find them any more useful in the Query Area, “*those correlations were in the other thing as well, and they weren’t useful for me in the insights*” (P10).

P#	Number of Views						Total
	Comparison	Correlations	Extremes	Outlier	Trends	Variability	
P1	1	4	2	-	-	-	7
P2	-	-	-	-	-	-	-
P3	2	3	3	-	1	-	9
P4	-	-	1	1	5	2	9
P5	14	10	-	-	21	3	48
P7	-	-	-	-	-	-	-
P8	-	2	-	-	-	-	2
P9	2	20	3	1	-	-	26
P10	-	1	-	-	1	-	2
Total	19	40	9	2	28	5	103

Table 6.10: Query Result Views By Query Type and Participant

Table 6.10 shows the number of query results that participants viewed, separated by query result type. These results suggest that when using this feature correlations tended to be of most interest, with the exception of P5 who appeared to prefer reviewing trends. P2 and P7 viewed no results.

Figure 6-10 shows a comparison of participants’ Query Area usage over time during their participation in the study. Similar to the Insight analysis, queries tend to be focused on one or two types of analysis, in a single session, with a few minor exceptions.

### 6.5.6 Utility of Queries

Participants’ comments with regards to the queries they conducted, centred around correlations. As is evidenced by the quantitative finding (see Section 6.5.5, Table 6.10) this was the most popular use for the Query Area. However, the number of participants

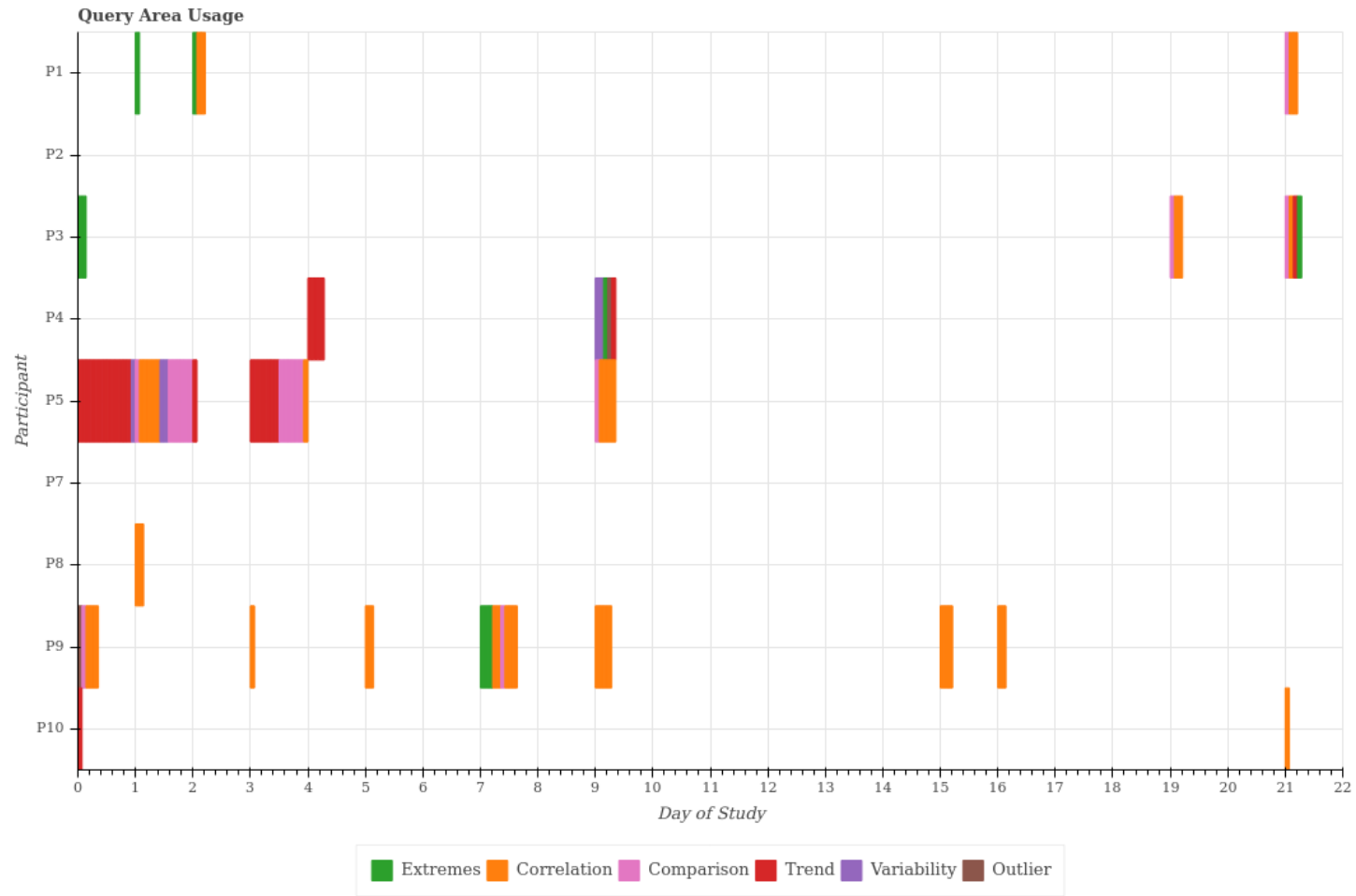


Figure 6-10: Query Area Usage By Participant

who spoke about the usefulness of specific query types was limited. P5 suggested that they tried a variety of queries but mostly stuck with correlations,

*I did try comparisons but I found that the graphs that generated were not as useful. So mostly did... And the trends a bit as well was useful but the correlations, I mostly generated the correlation (P5)*

Interesting to note, is that P5 described using the textual description of correlations and values rather than the graph when interpreting data, “*if I couldn’t make sense of the graph then I would just use the correlation value it gave you at the bottom*” (P5). P9 opted to make use of the correlation queries as it was the one the most easily understood, “*I used the correlation one because I found it most straightforward*” (P9). Both comments suggest that participants were driven to explore insights that provided simplicity in sensemaking, allowing them to understand their data in a way that they found easiest. These results suggest that users of analytical systems of this kind may shy away from functionality that provides analyses of data that they consider more difficult to interpret, both in terms of the data analysis method being performed and the way in which the results of the analyses are presented. Experienced members of the Quantified Self movement have been noted to struggle with data interpretation, often this is in relation to having too much information and a need to simplify the number of variables being analysed (Choe et al., 2014) and previous work by Rapp et al. also highlights the needs to provide both simplistic and more detailed views of data for those who want them (Rapp et al., 2018).

### 6.5.7 Challenges Using the Query Area

Participants who did make use of the Query Area said that they were hampered by the inability to select data based on particular activities that they carried out whilst recording data. This was also noted in relation to the Insight Feed. This frustration caused participants to stop using the Query Area once they realised that they would not be able to conduct the analysis in the way they wanted,

*I probably didn’t use it as much as I would have done, once I’d got more than a week in maybe, because I’d realised that the way that the data was doing that chronological order meant that the differences in the activities that I’d undertaken was sort of corrupting the correlation a little bit. (P1)*

Another participant felt that the Query Area was only useful “*to have fun searching for spurious correlations*” (P10). This highlights users’ awareness that aspects of the analysis that they wish to perform are often reliant on contextual information on

which to ground the data that is extracted for their analysis. However, it also highlights users' distrust in the correlations as a means of analysing data at all. Echoing P3's comment in relation to the Insight Feed (*"I wouldn't think, 'Ah that correlates, therefore..."* (P3)), some users have a keen awareness of the fallacy of believing that correlation implies causation. This participant further extended this by highlighting that they could potentially find correlations in anything if they looked hard enough. Contextual information may provide a basis on which to ground the correlation and lend itself to more realistic and truthful interpretations of data. Context-based analysis has been highlighted in several prior works, where participants have expressed a desire to be able to contextualise their experiences (Choe et al., 2014; Epstein et al., 2015; Raj, Lee, Garrity, & Newman, 2019).

One result that was also noted in comments about the Query Area was when asked if they were able to find the data they were looking for, the participant confirmed that they could but that the results *"weren't what I was anticipating"* (P7) and suggested that they did not use the result because of this. The comment relating to data not being as anticipated was also seen in relation to results in the Insight Feed and again this highlights an issue in regards to how willing users are to trust the data they are presented with. Previous work has noted similar effects in the automated analyses of presented results which users find counter-intuitive, or against their expectation, leading to mistrust in the data due to an inability to explain the results themselves (Raj et al., 2019).

### 6.5.8 Time to Sense-making

As an additional analysis we looked at *when* participants were attempting to make sense of the data they had been capturing. The purpose of this analysis was to understand how soon after recording data users tend to attempt to make sense of their recorded data and whether there were difference in time based on the type of analysis being performed, i.e. reviewing Session Overviews, the Insight Feed or the Query Area. Figure 6-11 shows Recording, Overview, Insight Feed and Query Area session usage by each participant over their participation in the study. As with the results in Section 6.5.1, we consider these sense-making sessions as any interactions with the Session Overview, Insight Feed or Query Area feature of the application within a 10 minute window of each other.

Figure 6-11: Recording vs. Overview vs. Insight Feed vs. Query Area Session



Table 6.11 shows the mean time per participant from finishing a recording (i.e. capturing data) to performing data analysis using either the Overview, Insight Feed or Query Area. This calculation only considers a recording session and the analysis immediately following it. Therefore, if a participant made five recording sessions before performing analysis, our analysis only accounts for the time difference between the fifth recording session and the start of the session in which they subsequently analyse their data. These values therefore assess how soon sense-making analyses were performed after a recording, with smaller values suggesting participants were engaging with analysis methods more immediately. The column labelled ‘% Matched’ shows the number of sense-making sessions that had a single preceding recording session. Lower values in this column suggest that participants made several recording sessions prior to attempting to analyse the data using the specific feature.

P#	Mean Time (hours) to					
	Overview Session	% Matched	Insight Session	% Matched	Query Session	% Matched
P1	0.0141	16.6667	18.0386	54.1667	0.6382	25.0000
P2	0.0928	19.0476	0.0530	4.7619	0.0380	9.5238
P3	1.7782	54.1667	10.7392	50.0000	7.9394	12.5000
P4	0.1980	36.3636	6.1813	27.2727	0.2009	3.0303
P5	0.3771	77.7778	14.2911	55.5556	0.8333	33.3333
P7	0.0477	14.2857	0.1240	7.1429	0.0360	7.1429
P8	0.7105	83.3333	5.7136	38.8889	0.0632	11.1111
P9	4.6570	18.5185	5.8894	37.0370	5.3789	29.6296
P10	10.0295	42.8571	15.3751	71.4286	16.2927	28.5714
M	1.9894	40.3352	8.4895	38.4727	3.4912	17.7603
SD	3.1713	25.0158	6.1307	21.0891	5.2586	10.6569

Table 6.11: Mean Time from Recording to Sense-making Session

These values suggest that the Session Overview was used more often after a recording session than the alternative methods of exploration, followed by the Insight Feed and the Query Area. Overview sessions and Query Area session tended to be used within a shorter time after having completed a recording. P10 appears to be an outlier, in particular for their time to sense-making sessions using the Session Overview and Query Area. When these are excluded from the results, time to analysis with the Session Overview is reduced to less than an hour ( $M = 0.9844$ ,  $SD = 1.4914$ ) and time to sense-making with the Query Area reduces to less than 2 hours ( $M = 1.8910$ ,  $SD = 2.8396$ ). This suggests that users tend to analyse their data in close succession to recording new data using direct analysis methods, such as the Session Overview or Query Area. A lack of immediate use of the Insight Feed suggests that users use automatically generated analyses at different times than more direct analysis methods.

## 6.6 Discussion

### 6.6.1 Designing for Exploration and Enquiry

This study aimed to understand whether providing users with a means of exploring their collected data via two features, an Insight Feed and Query Area, encouraged further exploration and enquiry into the recorded data. Further to this we sought to understand how such features may influence the way users engaged with exploration of their data. We designed the Insight Feed as a way of bringing attention to data by generating statistical insights based on automated statistical analyses of users' recorded data. We designed the Query Area as a way of providing users with a means of performing statistical analysis on their own data using a set of predefined query types.

Compared to the Session Overview, which users could use to review summarised views of the data they recorded on arranged by session, we find that participants were not necessarily more likely to use the provided Insight Feed and Query Area features. Participants used the Session Overview as frequently as the alternative features. Query Area results were viewed a similar number of times as Insight Feed details. There was a total of 114 Insights viewed in detail and a total of 103 Query results being viewed. This is compared to a total of 136 views of the application's general Overview. Based on this quantitative measure it would appear that there is no single preferred method for viewing data. Participants' interview responses suggest that both may have their place, depending on individuals preferences and what they are trying to achieve when exploring their data.

Summary level details that appeared in the Feed were sufficient to provide value (e.g. "You are quite likely (80.31%) to be more creative when your alpha values are higher.") as is evidenced by a large number of Insight Feed views being focused solely on the Feed itself. Our data suggests that participants did not often 'drill down' to explore these types of insights in further detail. This suggests that the providing insights as natural language text may be useful for exploring data without the need to provide deeper abilities to interrogate the data from within this section. Tollmar, Bentley, and Viedma (2012) similarly notes that although users viewed graphs, they found the natural language presentations, as a simple representation, of data were "valued and often more appreciated" than more complex visual representations (Tollmar et al., 2012).

We found that some participants tended to spend more time interacting with the Insight Feed and Query Area features than they did with the Session Overview. This suggests

that participants found using the Query Area and Insight Feed a more engaging means exploring their data. However, spending longer does not necessarily mean more time actively engaged. Participants may have spent more time using a particular feature because it was not as intuitive or it required more time to access the data they were looking for. Equally, the feature with the least amount of time spent may have been easier to use to find answers efficiently. The fallacy of using time data as a measure of engagement has been discussed previously in HCI literature (Doherty & Doherty, 2019), and we recognise therefore that this quantitative interpretation does not tell the whole picture. As our participants did not report any major usability issues we suggest that the increased time is more likely to reflect engagement in actively exploring their data.

The Query Area provided participants with a means to generate queries that would allow them to interrogate data that was of interest to them and in the way that they saw fit. The participants suggested this was a good way of getting answers to questions that they had already formulated. Therefore, when designing tools for those with a clear diagnostic need or objective a query-style tool is likely to be beneficial. Participants felt that the algorithmically curated Insight Feed was useful for bringing attention to insights that they may not have thought to look for themselves using tools such as the Query Area. The Insight Feed also appeared to act as a catalyst for further data exploration by some participants. Therefore, when designing data exploration features, designers should consider whether the likely user will have an idea of the analyses that they would like to perform or whether they are likely to benefit from data being brought to their attention. Based on the overlapping styles of tracking characterised by Rooksby et al. (2014) we suggest that users engaging in *directive* or *diagnostic* tracking styles may find the Query Area more useful due to their desire to achieve specific goals or to discover relationships between specific facets of their data. Those engaging in *documentary* and *fetishized* tracking may be more likely to find ‘interesting’ insights being brought to the fore useful, as they may not initially have specific questions that they wish to ask. Those engaging in *directive* or *diagnostic* tracking may also find benefit in the Insight Feed, but likely only for those insights which are directly related to the facets they are interested in.

Analysis of the ‘time to sense-making’ data suggests that users are more likely to use general Session Overviews and Query Areas to perform analysis soon after recording, than they are to visit their algorithmically generated Insight Feed. The time required to generate insights likely plays a factor in this. In the Cognizance Tracker app, the Insight Feed was updated with new insights on a daily basis. More immediate generation of new

insights would have perhaps seen increased usage of this feature. Our results suggest that the rhythm of the algorithmic insight generation was a determining factor for how frequently participants used this feature. PI system designers could consider that more frequent algorithmic insights may encourage more frequent interaction with the data. This could be considered as means of reducing information overload by providing users with more frequent, but smaller, sets of insights. The results suggests that even previously generated insights appearing within the Insight Feed are not reviewed at the same cadence as someone might perform their own analyses, using either overviews or customisable queries. The result of this study highlight that there can be large gaps in between recordings and data reflection and sense-making activities. We suggest therefore, that Session Overviews and Query Areas will be used for more immediate analysis of data and therefore should be designed in such a way to provide sufficient access to newly recorded data. Insight Feed-style features are more likely to be used as retrospective means of analysing data. Insights Feeds are likely to be more useful in providing information over longer time periods and bringing to the fore information that users may not see when reviewing data with the sense of immediacy that may be seen in results from a Session Overview or Query Area.

### **6.6.2 Designing to Uncover Insights - Volume & Variety**

Based on participants' responses, we consider that the number of generated insights was too high, leading to issues of information overload, such as participants not being aware that certain insights were even being generated. Information overload can have several negative impacts both from a HCI point-of-view, relevant in context of PI such as struggling to find interesting information (Bernstein et al., 2010), and disregarding information sources, such as a news feed, as being a reliable source of information (Koroleva et al., 2010). Information overload has been noted as tending to result in dissatisfaction with the service being used and may lead to abandonment (Eppler & Mengis, 2008). The inability for participants to effectively browse the entire Insight Feed and access the diverse variety of insight types that it contains appeared to lead users to feel that the application was only capable of generating specific types of insights. We therefore advise that PI system designers should take particular care when setting filters or thresholds that will determine the volume of algorithmically selected insights to be presented to users. Generating insights at less frequent interval(s) (e.g. weekly instead of daily) could be a simple method of reducing the number of insights generated. Alternatively, by increasing the level of statistical significance we might reduce the number of generated insights and also highlight only the items which have the strongest likelihood of being true. However, determining what an appropriate level is

still requires further consideration. Tentative correlations allow users to see emerging, if not statistically significant, correlations and allows them to see how these evolve over time. However, re-analysis of the correlation insights generated during this study and setting a significance level of  $p \leq 0.01$  still resulted between 334 and 740 correlation insights being generated ( $M=545.89$ ,  $SD=143.77$ ) per person over a period of 21 days. Similarly increasing the level of significance for trends to  $p \leq 0.01$  resulted in between 43 and 99 Trend Insights being generated ( $M=74.22$ ,  $SD=20.56$ ) per person over a 21 day period.

Given the large number of metrics being compared, the approach taken in this study possibly suffers from the multiple comparisons problem. That is, given the large number of variables being compared there is an increased likelihood of the variables being compared producing a significant result as a result of Type I error. Using the Bonferroni correction method (Armstrong, 2014) could attempt to account for this. For example, rather than presenting results that are considered significant using a  $p \leq 0.05$ , using Bonferroni correction would we only consider the test significant if  $p \leq 0.05/N$ , where  $N$  is the number of comparisons. Taking Correlation Insights as an example, 105 comparisons were being made. Using Bonferroni correction we could have considered a result significant if  $p \leq 0.002$  (given the initial tentative p-value level of 0.2), or even further restricted significance to  $p \leq 0.0004$ , using the more accepted initial p-value level of 0.05.

The Cognizance Tracker application included the ability for participants to rate insights using a thumbs up and thumbs down icon as a general response to whether they liked the insight, as well as fine-grained attributes such as how interesting they found the insight, how much they trusted a particular insight, how usefulness they found it and how understandable it was. However, this rating system was not used by most of the participants (thus we have not performed any in-depth analysis of this data). This style of user-provided rating system interaction is included in a variety of systems to improve individual recommendations, e.g. video recommendation (e.g. YouTube, Netflix), or social media platforms to personalise data shown or make recommendations (e.g. Facebook). Prior work has suggested that users may not fully understand how these recommendation systems work and which interactions influence recommendations (Alvarado, Heuer, Vanden Abeele, Breiter, & Verbert, 2020). In our study, the thumb ratings were used 82 times by 2 participants. One participant provided 66 ratings and the other provided 16. The fine-grained responses were completed for all of these ratings. This type of rating mechanism may have provided a basis for filtering insights based on the ratings for each dimension. However, we now suggest that it may

be difficult to obtain user-provided ratings as a means of classifying the interestingness of generated insights and therefore suggest that designers should not depend on this type of mechanism to gather such information. Alternative methods to explore include other personalisation methods such as users providing information as to their preferences when they first start using the PI system, or building a profile based on which insights users interact with most. The latter being more difficult when considering participants did not tend to drill-down into these details. Another consideration may be to explicitly call out the benefit of the functionality to the user, this was not done in this instance as there was no recommendation system built into the application to provide this benefit.

A more simplistic design consideration to reduce the number of insights might be to only consider an insight as new if the general description is considered different. For example, in this study a correlation insight was considered different if the correlation coefficient and/or the p-value had changed from the previously generated correlation insight variables. However, in future it might be better to only consider an insight different, or ‘new’, and therefore is an appropriate insight to include in the algorithmically curated feed if the underlying meaning of the provided insight has changed, i.e. whether something is or is not correlated, whether there is or is not a trend detected, rather than considering the strength of the underlying statistical analysis. For example, if the previously generated insight suggested that two variables were positively correlated, we would only consider an insight new if the correlation changed to a negative correlation or there was no longer a correlation detected. This type of change requires a trade-off between the detail with which a generated insight is presented and the volume of insights that are shown to the user, i.e. reducing granularity in order to mitigate information overload. Another alternative may be to use alternative sources of data as a filtering mechanisms to determine ‘interestingness’ of information, such as Google Trends data (e.g., S. L. Jones & Kelly, 2018).

A related factor that also influences the volume of insights presented within a system is the variety of insight types considered in the analysis. We note that in our analysis participants had marked differences in the insight types they considered helpful, useful and understandable. This variability in user’s preferences is also seen in prior work which highlights that the sensemaking process tends to vary depending on factors that can be personal to the individual who is analysing the data (Mamykina, Smaldone, & Bakken, 2015). Reducing the number of available insights could have a positive effect in reducing the volume of insights but could also have an effect of limiting the overall utility of the system.

### 6.6.3 Designing for the Sceptic

Previous work suggested that users found some statistically generated insights were obvious (e.g., S. Jones & Kelly, 2016; Bentley et al., 2013). However, some participants in our study showed a tendency to question how much they could trust the generated insights as it did not match their expectations. This could be due to the novelty of many of the metrics provided, thus participants were not aware of how the values are generated and secondly how they would behave in relation to other metrics. Additionally, participants were using a novel sensing technology and previous studies have highlighted increased levels of scepticism among what may be deemed the late adopters of technologies (H. C. Kim, 2015). We do not have data that specifically attempted to assess levels of trust or belief and willingness to adopt new technologies during this study. We therefore suggest that further research is required in this area to determine what role this may play but also suggest that PI system designers should bear in mind the levels of trust that users may be willing to assign to presented results. In Chapter 7, we present the results of a study intended to explore the factors that impact the formation of trust within the context of novel sensor data.

### 6.6.4 Designing for Context-based Analysis

One thing that was overlooked in the design of Cognizance Tracker was the ability to specifically filter and compare data between sessions based on context, e.g. activity, location. Previous work has suggested that a lack of contextual information can lead to users not being able to gain insights from their data (Choe et al., 2014). Raj et al. suggests that contextual data such as notes, location and mood provide anchors on which analyses can be framed, and aided in finding and explaining context-based trends, discovering factors, context-specific actions, as well as hypothesising about alternative context-based outcomes (Raj et al., 2019). The desire for contextual information is consistent with comments from our previous studies (see Chapter 3 & 4). This type of comparison appears to be of interest in both the Insight Feed and the Query Area. Having included the ability for participants to tag data with activities within the application, it was an oversight to only make this information available in the Session Overview section and not in the Query Area or the Insight Feed. As a design consideration for future implementations of similar features we suggest that designers might consider including the possibility for users to run automated insight generation based on user-specified subsets of data, such as activity. This may have a limitation that it would still require user input to determine what is a relevant subset of data to run the analyses on, and thus reduces the level of automation that can be provided. However,

it is possible that insights could be generated for all subsets of a particular activity and then only retrieved when requested by a user.

## 6.7 Conclusion

In this chapter we presented the results of an analysis of two features, an Insight Feed and Query Area, which aimed to encourage exploration and enquiry of data collected using a multifaceted PI tool during a 21 day in-the-wild study with 9 participants. We find that there were differences in time to sense-making using different features with users more likely to use Query-based analysis soon after a recording sessions. Insight-based analyses were not likely to be performed in close succession with recording sessions. However, Insights provided people with no initial questions about their data ideas that they could then explore further using queries, suggesting that generated insights can be used to encourage exploration and enquiry into PI data. Through our analysis of quantitative data from recorded application, interactions and qualitative analysis of interviews, we present four design considerations and suggestions for future research directions:

- Designing for Exploration and Enquiry
- Designing to Uncover Insights
- Designing for The Sceptic
- Designing for Context-based Analysis.

We reveal that feature usage may differ based on differences in users intended tracking style i.e. what and how they are seeking to analyse their data and whether they have initial questions that they are seeking answers to. Design considerations also vary as to whether the user is likely to benefit from immediate analysis of the data or whether they may benefit from insights being brought to the fore over a longer time period, as and when a relevant insight has been generated. We suggest that insights should be focused on providing quality information rather than larger quantities of data that may not be of benefit to the user. Additionally, we suggest that a significant focus should be given to context-based analysis rather than just general analysis of data over time. Possible future research includes further understanding the ease at which these features enable users to gain understanding of their data. Although engagement is a possible metric for users' interest in using certain types of feature, another metric that could be investigated is whether they are able to achieve a better understanding of the data they are analysing. This was not explored in this study, as our questions largely



focused on utility as a method of exploring and bringing data to attention, rather than comparative usefulness. Another aspect that requires more research is how users come to trust and believe the data they are presented with, noting that some participants had a tendency to not trust the generated insights they were shown. We explore this latter question in the following chapter.

## Chapter 7

# Trust in Novel Sensing Technologies

## 7.1 Chapter Overview

The study described in this chapter aims to better understand how people calibrate trust in data visualisations of their own data from novel sensing devices. Our previous qualitative research suggests that people are often inclined to trust data produced by devices and visualised for them, despite varying degrees of accuracy and reliability in consumer grade sensing devices. This chapter seeks to better understand the role that the visualisation of recorded data plays in shaping a user's trust in the sensor and the data that it provides.

O'Hara (2012) describes trust as a consideration by a person of the trustworthiness of an agent (someone or something). The concept of trustworthiness of an agent is the characterisation that, all things being equal, the agent will do what it claims it will do. Thus we consider trust as an attitude of a person toward something, or someone, with the property of trustworthiness. Trust is considered to be an attitude towards the trustworthiness of another (O'Hara, 2012).

In their analysis of how trust relations are formed with IT artefacts, Söllner, Hoffmann, Hoffmann, Wacker, and Leimeister (2014) suggests that although previous research has considered IT artefacts as playing a mediator role in between-human trust, people also place direct trust in IT artefacts due to their increasing role in providing help to the user to achieve specific goals. The authors also highlight the importance of researching initial trust impressions, citing previous research by Koufaris and Hampton-Sosa (2004) and Harrison McKnight, Choudhury, and Kacmar (2002) suggesting that initial trust forms the basis on which future interactions will take place.

PI data may be used by the collector to affect change about themselves in their daily lives. Given the variability in accuracy of such technologies it may not always be clear to the user how reliable and trustworthy this data is. In our previous studies we found that some people had a tendency to trust the data provided to them due to inherent trust in technology with little thought for how trustworthy the devices are, whilst others tended to determine trust through research and public reputation. A consideration based on the results of Chapter 5 is that the perceived accuracy of a device may influence trust, with users suggesting they would trust a device more if it was more accurate. Based on the results in Chapter 5 and Chapter 6, we suggest that a users' expectations of how data should behave may affect their willingness to accept results that are presented to them. Participants suggested a tendency to trust the data that met their expectations and question those which do not.

An understanding of the factors that influence perceptions of trust in PI systems will enable designers and developers to appropriately design applications that enable realistic expectations from users about the reliability of the data being collected and analysed, and confidence that should be placed in this data.

We present a study that seeks to understand what factors affect the trust that people place in sensing technologies based on visualisations of data collected by the technology. Once again we focus on the context of gathering and analysing personal informatics data for self-reflection and self-understanding. In particular, we investigate how people form an assessment of trust in novel sensors (e.g. cognitive/EEG sensors such as the NeuroSky MindWave) and how that assessment may differ from more established and familiar sensing devices (e.g. heart rate trackers). This comparison will allow us to better understand the factors that may affect people when using new sensing technologies, in particular when determining whether to trust the data presented by the device.

In this study the following hypotheses were tested:

- H1** There will be a significant increase in user trust in data when the data more closely reflects users' prior predictions about data values.

Confirmation bias, or the human tendency to interpret data in a way that confirms our own initial hypotheses or expectations (Nickerson, 1998), may be a factor that influences trust. Kersten-van Dijk, Westerink, Beute, and IJsselsteijn (2017) posits that confirmation bias may be a factor that influences people's ability to gain insights through self-tracking. This study will seek to understand whether those biases based on visualised data are also a factor in initial trust, which could affect the ability to gain insights from PI data. In assessing the influence of social information on peoples' interpretation of data, Y. S. Kim, Reinecke, and Hullman (2018) asked participant to draw expectation for data trends and true data trends were revealed afterwards. This post-hoc revelation also included other peoples' expectations for the data that were either congruent or incongruent with the participant's expectations. The authors found that "[i]f a participant and other people disagreed with the trend of the base data, the participant was less likely to trust its accuracy, was more likely to stick with their initial expectations, and their updated expectations were less aligned with the trend of the base data" (Y. S. Kim et al., 2018). The authors suggest that this social influence acts as a factor to strengthen confirmation biases. Thus we hypothesise that people may be less likely to question the reliability and accuracy of sensing technologies when the data produced aligns more closely with their prior expectations.

- H2** There will be a significant difference in user trust in a personal informatics device

and the data it produces according to the precision at which it is presented within a visualisation. Therefore, rounded values will be less trusted than precise values.

The effect precision of data has on people’s interpretations of that data has been explored in a variety of disciplines. In assessing precision as a cue to perceptions of confidence, Jerez-Fernandez, Angulo, and Oppenheimer (2014) found that when presented with measurements, participants rated their perceived confidence in the person who recorded that values more highly when the values included more significant figures (e.g. 1234 vs. 1200). On the other hand, Olsen (2018) found that when people were asked to assess trust based on data visualisations of political data, that they were no more likely to trust data that was more precisely presented than data that was rounded. Our hypothesis, therefore, seeks to determine the effect of precision of visualised data in relation to personal data presented within the context of PI.

**H3** Trust in the sensor data will differ based on the level of knowledge the participant has about how a particular sensing technology is measuring and capturing data.

We hypothesis that participants’ trust may be influenced by their ability to understand how data is collected by the device. In turn, we expect that knowledge of how data is collected will correlate with familiarity of a particular device. Hence, we expect that overall trust in the NeuroSky device and metrics would be lower than the Fitbit device for most participants due to expected limited experience with EEG data. However, this may be affected by other factors. In line with prior work, we expect that a person’s trust in a personal informatics device may be inversely related to their knowledge about the machine (Muir, 1987).

**H4** Trust will differ according to whether the data is presented synchronously (in real-time) vs. asynchronously (post-task), based on participants’ ability to check ‘in the moment’ whether the data they are seeing corroborates their expectations.

In our previous studies (see Chapter 3 & 4) we note that participants found the inclusion of live views to be a distraction from the recording process. However, this view is often included in PI tools and therefore its effect on trust is still worth considering. Having the ability to see the data being recorded in real-time may influence people’s perceptions of the data. Participant’s may be influenced by the ability to correlate lived experiences with the data in real-time, i.e. does the data I am seeing now seem accurate/trustworthy based on my current feelings?

In comparison to participants who are required to rely on their memory for post-hoc evaluation, participants who are able to see the data in real-time may be influenced by

this. Thus, live views of data could influence recall biases, a systematic error in person's ability to correctly recall previous events (Spencer, Brassey, & Mahtani, 2017), thus affecting their levels of trust.

Irregardless of whether the data matches their prior expectation, people may come to believe that the measurement is a true reflection of what was recorded. Data shown post-hoc may then become considered a true reflection of the data being by virtue of having seen it progress. This may be another instance that is seen in the previously mentioned research whereby people will defer to the algorithm despite believing otherwise (Hollis et al., 2018) or due to increased belief in the technology due to lack of knowledge (Muir, 1987).

The study described in this chapter seeks to understand whether the factors that influence trust and confidence in data visualisations are also seen as factors that influence the confidence in Personal Informatics technologies. This chapter provides an understanding of the effect of data visualisation on initial formations of trust within the context of PI.

## 7.2 Methodology

In order to answer our hypotheses we wanted to observe participants in a situation where they would experience fluctuations in attention, meditation and heart rate in response to a particular stimulus which could be measured using wearable consumer-grade sensing devices. This would enable us to gather data which could be visualised for the participants in a way that they would be subjectively aware of how they felt in comparison to how the data suggests they may have felt during the experiment. This presents a situation similar to that of a PI system user; recording data in relation to specific tasks and then assessing that data.

Thus, we designed a lab-based study where participants were asked to perform a stress-inducing computer-based test, called a Stroop test (MacLeod, 1991), whilst wearing a NeuroSky headset and Fitbit Versa smartwatch. Study participation took less than one hour per participant. The study was run over a period of one week. We chose the Stroop test as it has previously been seen to have a measurable effect on physiological responses, acting as a mild stressor, and increasing heart rate (Renaud & Blondin, 1997). Previous work has also seen an effect on Meditation and Attention values during the Stroop test, when these values are measured using a NeuroSky headset (Crowley et al., 2010).

During a Stroop test participants are presented with a colour, as a word, for example the word 'RED'. However, the colour of the font in which that word is presented may not be congruent with the word. For example, the participants may be presented with the word 'RED' in a green-coloured font, or the word 'GREEN' may be presented in a yellow-coloured font, and so on. The participant's objective during the Stroop test is to correctly respond with the colour of the font presented on-screen, rather than the word displayed on screen. As the test, and trials, progress the time between presented stimuli is reduced making the task harder for the participant.

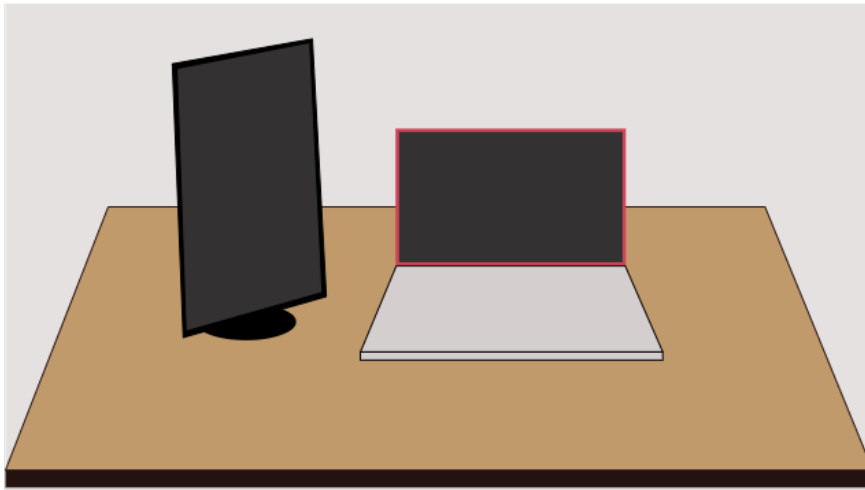


Figure 7-1: Study 4 - Experimental Setup

As shown in Figure 7-1, participants were seated at a desk with a laptop and an additional 24" monitor. The monitor was placed to the left of the laptop, orientated to a portrait position. Before beginning the test participants were provided with an information sheet and consent form to sign (see Appendix E) and any questions they had were answered.

After this the devices were introduced to the participant so they were aware of what each device is called and how it is worn. The NeuroSky and Fitbit Versa were then placed on the participants' head and wrist respectively, by the researcher. Participants were then asked to provide demographic information and complete the pre-experiment survey via the laptop, survey questions were displayed on the additional monitor. Demographic data included the participant's age, gender, educational background, previous experience with personal informatics and previous experience with brain-computer interfaces.

The study methodology received ethical approval from the university's Psychology Research Ethics Committee (PREC) (Reference: 20-011). Data pertaining to this

study is stored in accordance with the data management plan in Appendix F.

### 7.2.1 Survey & Stroop Test

The pre-experiment survey provided participants with the manufacturer’s description of the device and the manufacturer’s description of the value it was recording, for example Fitbit was described as follows, based on the Fitbit website:

#### **Fitbit Versa**

Live your best life with Fitbit Versa family-health and fitness smartwatches that last 4+ days and feature activity, sleep and heart rate tracking, notifications, apps, & more.

#### **PurePulse**

Photoplethysmography is a long word with a short explanation: light can be used to measure blood flow.

When your heart beats, blood flows, and the volume of the blood in your wrist changes. Blood—interestingly enough—absorbs green light. The higher your blood volume is, the more green light is absorbed.

To calculate blood flow, PurePulse shines a green light onto the skin and uses light detectors called photodiodes to measure how much light is being absorbed. This measurement is used to determine how many times your heart beats per minute.

Participants were then asked to respond to how knowledgeable they were about the device and how confident they were that the devices would reflect the values being recorded, on scales from 1 to 10. Open-ended responses were also recorded for *why* they provided those ratings:

- How would you rate your knowledge of Fitbit Versa device?
- How confident are you that this device will reflect your heart rate?
- Why?

A Stroop test task is made up of 5 trials. Participants were allowed to get comfortable and were asked to press the spacebar to begin when they were ready. During each trial participants were presented with 16 different word-colour combinations, displayed on the laptop display. Participants were requested to respond by pressing the R, G, Y, B keys on the keyboard to the corresponding colours red, green, yellow and blue



respectively. Within each trial the time a participant has to respond to the word-colour combination is consistent but as the trials progress this time is reduced. In the first trial there was up to 2.5 seconds to respond to each word-colour combination before the next word-colour combination is displayed. In the final trial the participant has 0.5 seconds to respond to each combination. If the participants responded correctly or incorrectly the next combination was shown immediately.

Participants performed one Stroop test task to demonstrate how the trials progress during the task. Based on this demonstration run the participants were asked to draw a line on graph (see Figure 7-2), via a tablet, of how they expected the values for heart rate, Attention and Meditation to change as they progressed through a second Stroop test task. Similar to the approach of drawing expectations seen in Y. S. Kim et al. (2018). The graph had possible values on the y-axis and vertical lines representing the end of each trial, labelled 1 to 5.

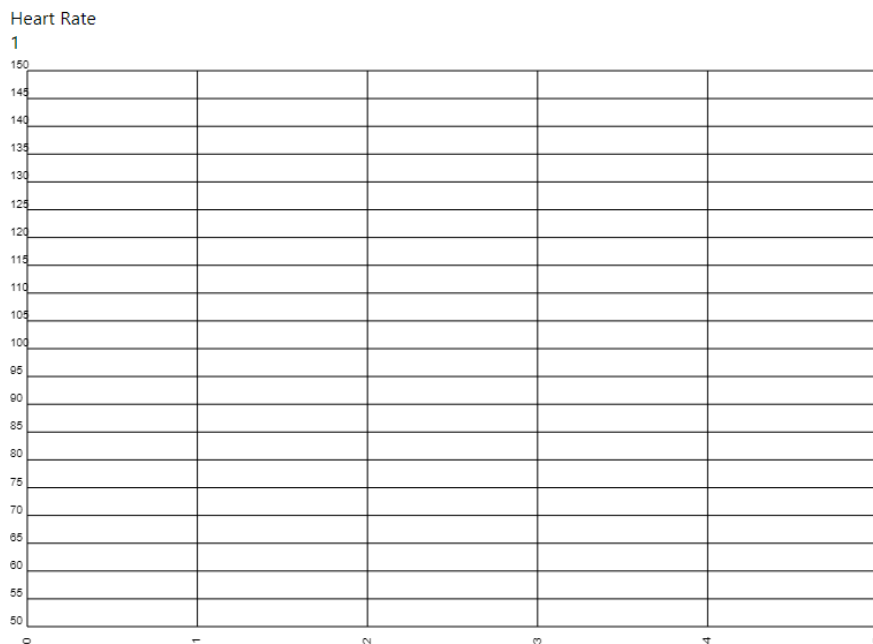


Figure 7-2: Heart Rate - Participant Expected Trend Graph

The participants then performed the second Stroop test task (referred to as the ‘main Stroop test’). After this participants were presented with two pages of scaled questions in response to data visualisations. The first page contained the data visualisation of heart rate values. The second page contained the data visualisation for Attention and Meditation values. For all graphs the y-axis presented the value of the recorded data and the x-axis presented time (up to 300 seconds/5 minutes). Additionally, vertical lines

were added indicating the end of each trial. Below each of these graphs participants were asked to give Likert scale responses for each metric:

- How confident are you that the data shown here accurately reflects your <heart rate / attention / meditation> during the task?
- How closely does the data shown here reflect your prior expectation about how your <heart rate / attention / meditation> might change during the task?

Each page also asked once per device:

- How confident are you that the device used to collect this data is trustworthy?

The survey pages as presented to the participants can be see in Appendix E.3.

The data visualisation that the participants were presented with at the end of the experiment was dependent on which one of the four study conditions they had been assigned to (as seen in Figure 7-3).

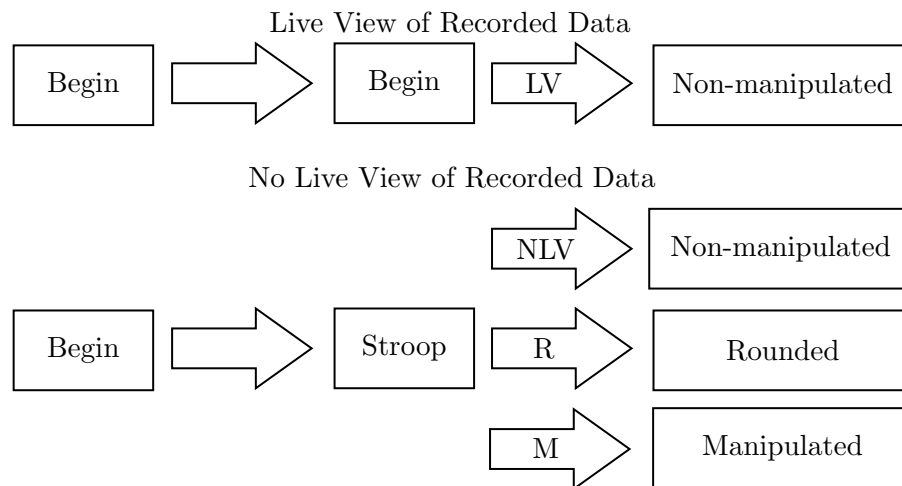


Figure 7-3: Four visualisation conditions

In Condition LV (Live) participants were able to see a live view of the data being recorded as they performed the main Stroop test, on the additional monitor. During a Stroop task, between each trial participants had an 8 second pause during which they were advised to ‘Check your readings’. As part of evaluating **H4**, this pause gave participants time to see their data as it was recorded. Without this pause it was felt that there would be no time for a participant to assess the data while concentrating on the Stroop test task. A beep sound was played 1 second before the next trial began to draw participants’ attention back to the task from the data visualisation. In the post-

experiment questions, participants in this condition were presented with visualisations of the actual data that was recorded by the sensors during the main Stroop test.

In Conditions NLV (Non-Live), R (Rounded) and M (Manipulated) participants were unable to see the data that was being recorded in real-time as they performed the main Stroop test. In these conditions there was still a period of 8 seconds and a beep to alert the participant that the next trial was starting. Participants in these conditions were not prompted to check their real-time data recordings and were presented with a blank page on the secondary display.

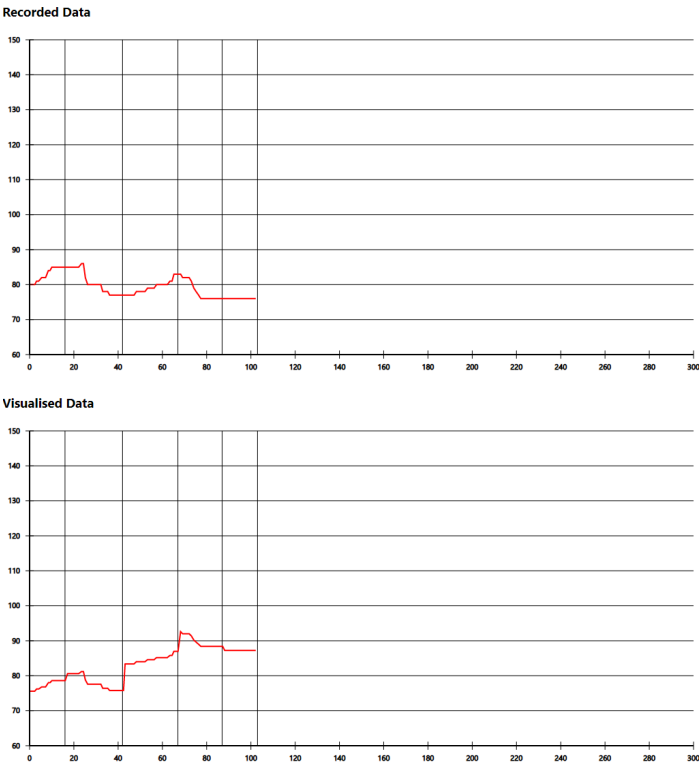
In the post-study questions for participants in Condition NLV the data presented was not manipulated in any way, as in Condition LV. In Condition R the visualised data was rounded to be less precise, i.e. all values were rounded up or down to the nearest value of 10. In Condition M the data was manipulated such that the presented data takes on characteristics of the participant's expectation drawn prior to undertaking the second Stroop test. In Condition M, we chose to perform manipulation of the data that was presented to participants to test the mediating effect of the relationship between expectations and match on trust. We do not expect, or suggest, that sensing devices should or do manipulate data, rather we performed this experimental manipulation to test the mediating effect of the relationship between expectations and match, on trust.

### **7.2.2 Visualisation Manipulation**

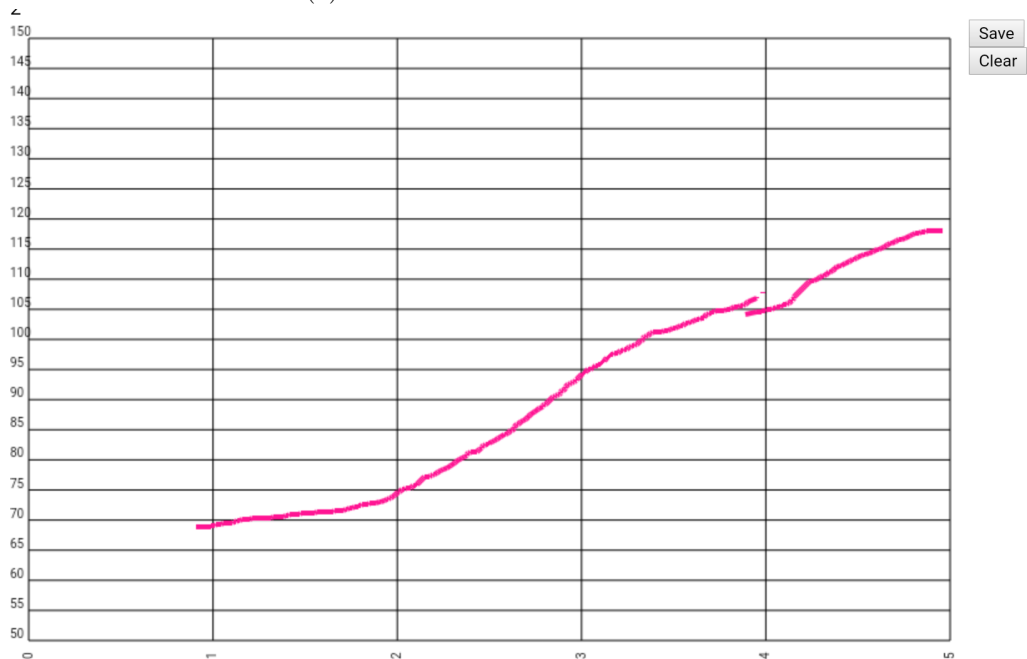
To perform the manipulation of the data in Condition M the values are taken from the drawing participants were asked to draw on a touchscreen tablet, as described in the previous section. The drawn values were combined with the sensor's recorded values within the same trial period, using a weighted average. The values presented to participants in Condition M visualisations used a weighted average such that there is a slight artificial skew towards the expected data values, rather than the true data. As in all other conditions, participants were asked to rate (on a scale from 1 to 10) how confident they were that the presented data accurately reflected the metric being recorded, how closely the presented data reflected their prior expectation, how trustworthy they believed the particular device is.

Figures 7-4, 7-5 and 7-6 show examples of the data that was recorded, the data that was displayed to the participant, and the participant's drawing of their expectations, prior to the main Stroop test task, in Condition M. The graphs labelled as 'Recorded Data' are the true values that were recorded from the device during the main Stroop test task. The graphs labelled as 'Visualised Data' are the graphs that were shown

Heart Rate



(a) Heart Rate Data Visualisation - P9

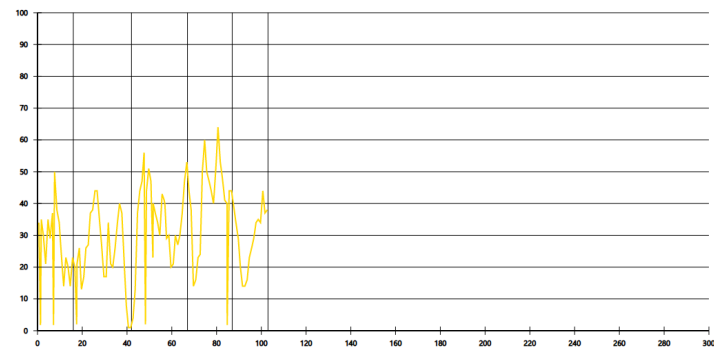


(b) Heart Rate Expected Values - P9

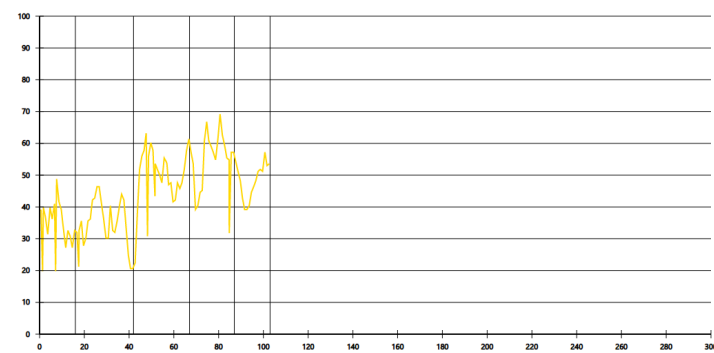
Figure 7-4: Data Manipulation Example Graphs - Heart Rate

## Attention

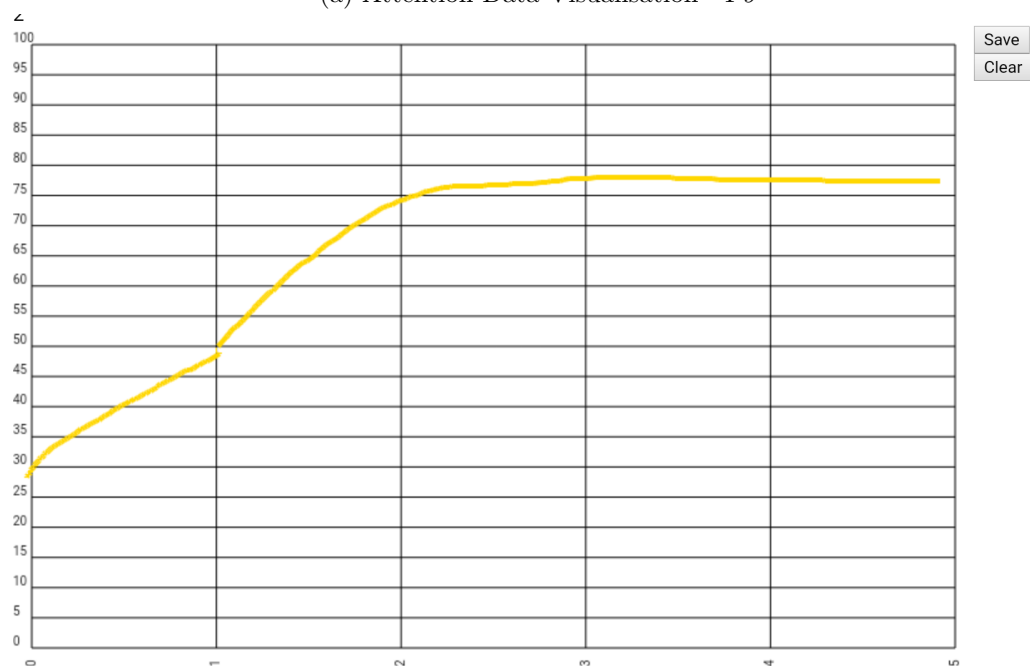
Recorded Data



Visualised Data



(a) Attention Data Visualisation - P9

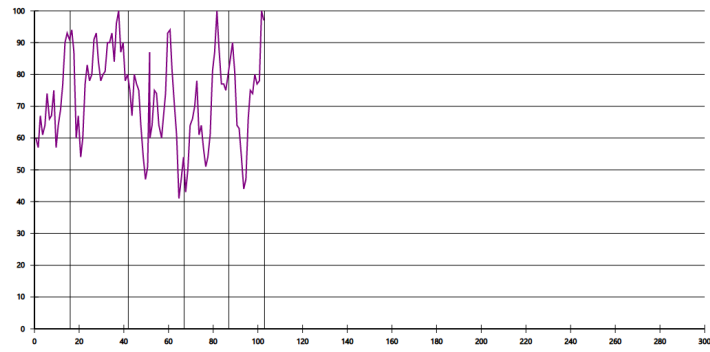


(b) Attention Expected Values - P9

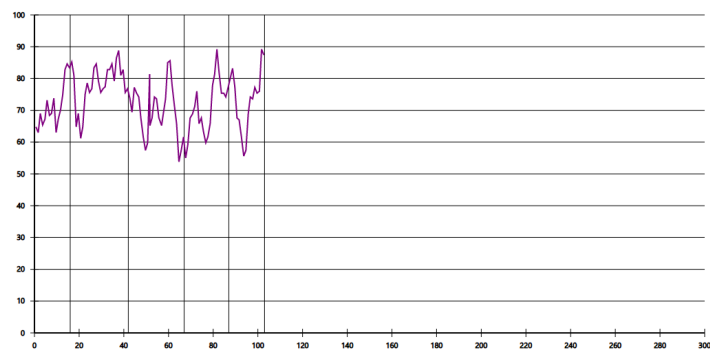
Figure 7-5: Data Manipulation Example Graphs - Attention

## Meditation

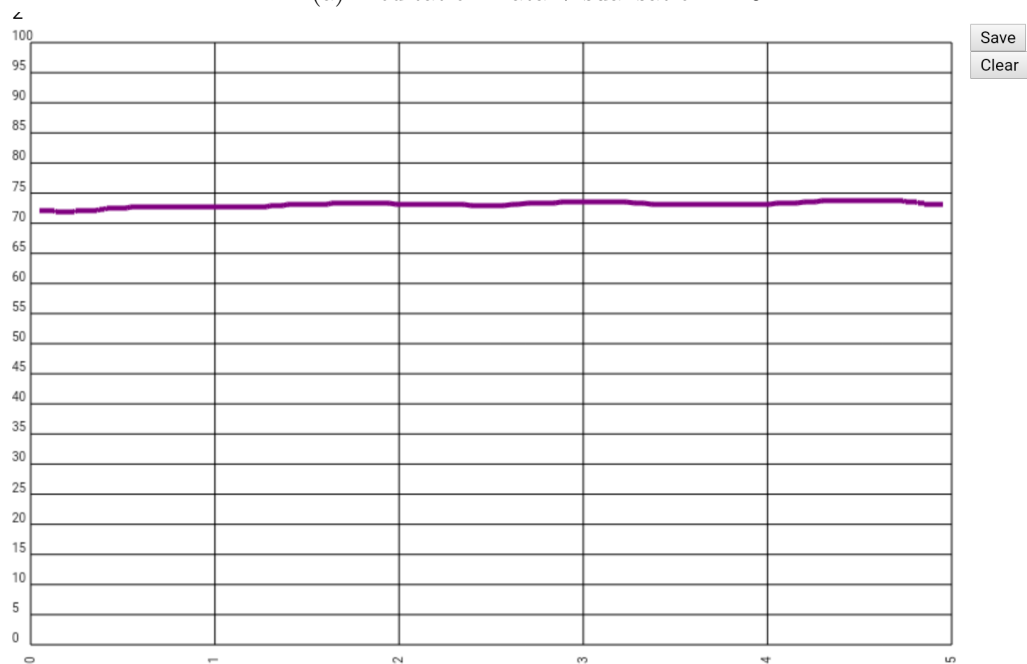
Recorded Data



Visualised Data



(a) Meditation Data Visualisation - P9



(b) Meditation Expected Values - P9

Figure 7-6: Data Manipulation Example Graphs - Meditation

to the participants when asked to answer the post-study survey questions. The final graph is the participant’s drawing of how they expected their values to change over time. The ‘Visualised Data’ graph therefore shows a weighted average between the true values and the drawn values.

The effect of the artificial manipulation is most pronounced for the heart rate graphs in this example. The data shown to the participant shows an increase in heart rate toward the end of the trials, whereas the recorded data actually showed a decrease in their heart rate (see Figure 7-4a). This is due to the upward trend they have drawn, based on their expectations (see Figure 7-4b). In the Meditation and Attention examples the effect is less pronounced. However, for Attention the visualised values shift slightly up from the true values (see Figure 7-5a), whilst the visualised Meditation value show a decreased overall range (see Figure 7-6a), due to the effect of averaging around an almost constant value, expected by the participant.

---

**Algorithm 1** Data Manipulation Algorithm

---

```

1: function MANIPULATE(drawnData[ ], realData[ ])
2:   augmented[ ]  $\leftarrow$  [ ]
3:   for  $i \leftarrow 1$  to num_trials do
4:     possible[ ]  $\leftarrow$  drawnData where drawnData.trial =  $i$ 
5:     recorded[ ]  $\leftarrow$  realData where realData.trial =  $i$ 
6:     for  $j \leftarrow 0$  to length(recording) do
7:       index  $\leftarrow j / \text{length}(\text{possible})$ 
8:       if index < length(possible) then
9:         aug_val  $\leftarrow$  recorded[ $j$ ] * 0.6 + possible[index] * 0.4
10:      else
11:        aug_val  $\leftarrow$  recorded[ $j$ ]
12:      end if
13:      augmented.append(aug_val)
14:    end for
15:  end for
16:  return augmented
17: end function

```

---

The algorithm used to determine the graphed values can be seen in Algorithm 1. This is a weighted average of  $0.6 * \text{real\_value} + 0.4 * \text{drawn\_value}$ . Drawn values are pre-converted to second intervals to account for the difference between pixel density on the touchscreen tablet and the recorded data being captured at rate of 1 value per second. Extraction of the drawn value is such that if a corresponding drawn value is not available (i.e. a participant did not draw any data for a particular trial) the ‘real’ data is used with no manipulation, to prevent automatically decreasing values by

60%. The data was manipulated using this formula following a trial and error process, adapting the weights of each component, whilst seeking to achieve a balance between realistic looking data that preserved salient features of the real values, such as the noisiness of the EEG data, but also shifting the data far enough towards participants' drawn expectations.

### **7.2.3 Participant Deception Considerations**

In designing this experiment consideration was given to whether it would be possible to test our hypotheses in a manner that did not require deceiving participants.

One alternative that was considered was showing a manipulated and non-manipulated data visualisation and asking the participant to choose which one they thought was genuine prior to rating their trust in the device. Another considered alternative was to manipulate participants' initial expectations of how the data should look such as telling them in what way the data may behave (e.g. The data from the EEG headset will be noisy).

However, in both of these situations it was felt that they would bias the participants expectations in such a manner that would bias their responses. If the participants were aware that one data visualisation was true and one was not a true representation then this may influence responses and expectation. Similarly it was felt that any expectation management prior to the study would alter the participants' responses to not be a true reflection of their own expectation of the devices and the data produced.

The deception was revealed to participants after completing the second set of survey questions. Participants in Condition M were then shown the true data captured by the devices alongside the manipulated data. Additionally, participants were given a second opportunity to consent or withdraw from their participation after the deception was revealed.

### **7.2.4 Participants**

Participants were recruited via University mailing lists, online and physical noticeboards, and word of mouth. Participants were entered into a draw to win 1 of 3 £20 Amazon gift vouchers for participating in the study. Any participant aged over 18 was eligible to participate in the study. The exclusion criteria was left open to enable participation from a broad, general sample of the population.

42 participants took part in the study, 6 participants' results were excluded from ana-



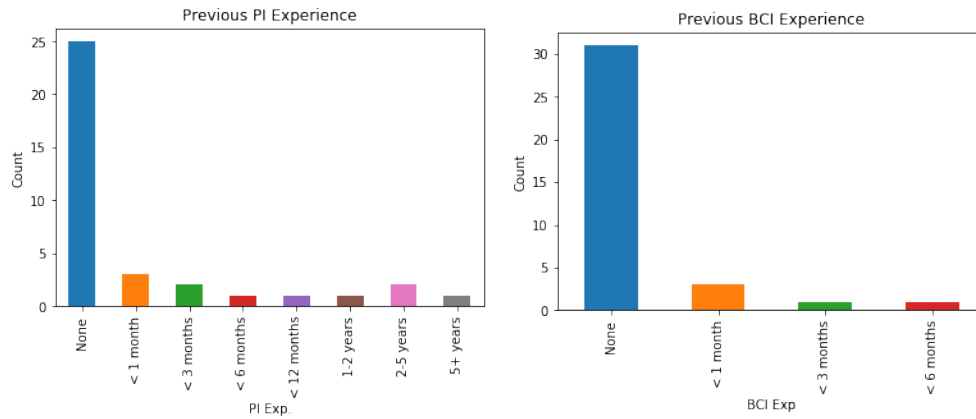


Figure 7-7: Previous BCI and PI experience

lysis due to issues with data collection resulting in either no data or only partial data being collected. For two participants survey responses were not recorded by the software. For the other four participants issues with recording of sensor data resulted in no data being visualised for them to base responses on during the post-study survey. Of the 36 participants' (F=17, M=19) data that is included in the analyses discussed in the following sections, participants' ages ranged from 18 to 53 (M=27.47, SD=6.19).

The majority of participants had no previous experience with BCI technologies (86.1%) or previous experience with personal informatics (69.4%). All participants had less than 12 months experience with BCI technologies, and the most experience with PI was 5+ years (see Figure 7-7).

Participants were assigned to a condition using block randomisation (Suresh, 2011). Participants picked a time to attend the lab to perform the tests based on their own availability and were assigned to the next available condition. The final number of participants per condition were as follows: LV=8, M=9, NLV=8, R=11.

## 7.3 Results

The following subsections presents the results of our quantitative analysis of participant responses. In addition we present our qualitative analysis of participant's open-ended responses.

Age	Gender	Highest Level Of Education	Experience with PI	Experience with BCI
33	F	University Higher Degree (e.g. MSc, PhD)	< 1 month	< 1 month
24	F	University Higher Degree (e.g. MSc, PhD)	< 1 month	< 1 month
28	M	University Higher Degree (e.g. MSc, PhD)	< 1 month	< 1 month
33	M	University Higher Degree (e.g. MSc, PhD)	2-5 years	< 1 month
20	M	First degree level qualification	< 1 month	< 3 months
23	F	First degree level qualification	None	< 6 months
39	M	First degree level qualification	None	1-2 years
33	M	First degree level qualification	< 12 months	None
25	F	University Higher Degree (e.g. MSc, PhD)	< 3 months	None
29	F	University Higher Degree (e.g. MSc, PhD)	< 3 months	None
24	F	University Higher Degree (e.g. MSc, PhD)	< 3 months	None
25	M	First degree level qualification	< 6 months	None
19	M	A/AS Level (or equivalent)	1-2 years	None
27	F	University Higher Degree (e.g. MSc, PhD)	2-5 years	None
31	F	University Higher Degree (e.g. MSc, PhD)	2-5 years	None
18	M	Other school	5+ years	None
24	M	First degree level qualification	None	None
22	M	First degree level qualification	None	None
25	M	First degree level qualification	None	None
22	F	First degree level qualification	None	None
24	M	First degree level qualification	None	None
32	M	University Higher Degree (e.g. MSc, PhD)	None	None
24	F	University Higher Degree (e.g. MSc, PhD)	None	None
30	F	University Higher Degree (e.g. MSc, PhD)	None	None
26	F	University Higher Degree (e.g. MSc, PhD)	None	None
25	M	University Higher Degree (e.g. MSc, PhD)	None	None
25	M	University Higher Degree (e.g. MSc, PhD)	None	None
34	F	University Higher Degree (e.g. MSc, PhD)	None	None
24	F	University Higher Degree (e.g. MSc, PhD)	None	None
32	M	University Higher Degree (e.g. MSc, PhD)	None	None
27	M	University Higher Degree (e.g. MSc, PhD)	None	None
29	M	University Higher Degree (e.g. MSc, PhD)	None	None
23	F	University Higher Degree (e.g. MSc, PhD)	None	None
22	F	University Higher Degree (e.g. MSc, PhD)	None	None
25	M	University Higher Degree (e.g. MSc, PhD)	None	None
53	M	University Higher Degree (e.g. MSc, PhD)	None	None
30	M	University Higher Degree (e.g. MSc, PhD)	None	None
23	F	University Higher Degree (e.g. MSc, PhD)	None	None
25	F	University Higher Degree (e.g. MSc, PhD)	None	None

Table 7.1: Study 4 - Participant Demographics

### 7.3.1 Confidence in Accuracy of Data Visualisation

Participants in each condition were asked to rate how confident they were that the data presented in the visualisation accurately reflected the values that the device is supposed to measure.

**How confident are you that the data shown here accurately reflects your heart rate during the task?**

Regarding the Fitbit and heart rate recording, participants in the manipulated condition rated their confidence that the visualisation accurately reflected their heart rate higher than in any of the other groups. Additionally, in the rounded condition there is a decrease in value ratings, suggesting that precision of values may influence confidence in a devices accuracy.

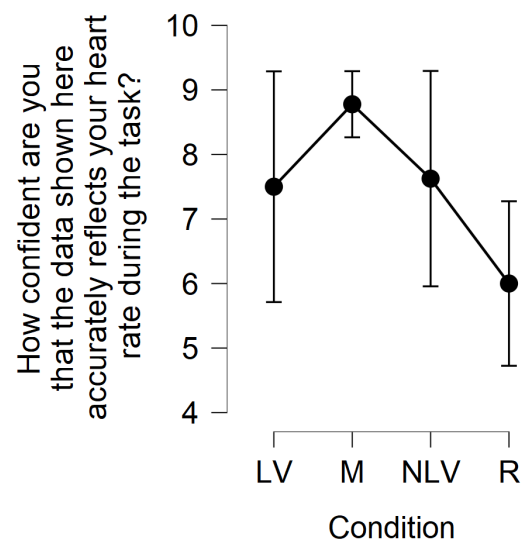


Figure 7-8: Confidence in Accuracy of Data Visualisation - Heart Rate

An analysis of variance showed that the effect of the data visualisation on confidence that the data shown accurately reflected heart rate was significant,  $F(3, 32) = 4.197$ ,  $p = 0.013$ . The mean range of these responses can be seen in Figure 7-8.

Post-hoc analysis with the Dunn test (used because of its non-parametric approach and due to the small number of comparisons<sup>1</sup>) indicates that there is a significant effect of the data visualisation on user's confidence that the data shown accurately reflected heart rate, between Condition M and R ( $z = 3.146$ ,  $p < 0.001$ ), as well as between

<sup>1</sup><https://www.statology.org/dunns-test/>

Condition NLV and R ( $z = 1.782, p = 0.037$ ). There were no other significant effects measured between other conditions.

**How confident are you that the data shown here accurately reflects your level of Attention during the task?**

An analysis of variance showed that the effect of the data visualisation on confidence that the data accurately reflected Attention during the task was not significant,  $F(3, 32) = 0.621, p = 0.607$ . The mean and range of the responses for this question can be seen in Figure 7-9.

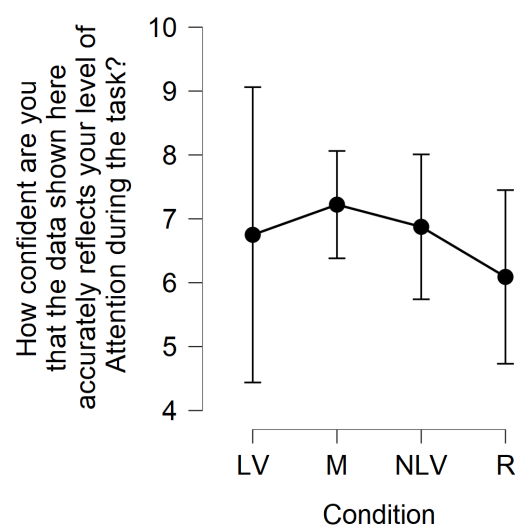


Figure 7-9: Confidence in Accuracy of Data Visualisation - Attention

All values appear to have similar means and the range of values for Condition NLV, R, and M. All fall within the range of rating responses in condition LV.

**How confident are you that the data shown here accurately reflects your level of Meditation during the task?**

For the visualisation of Meditation values, similar to what was seen for Attention, there does not appear to be a significant difference in confidence levels between conditions.

An analysis of variance showed that the effect of the data visualisation on confidence that data shown accurately reflected Attention during the task was not significant,  $F(3, 32) = 1.950, p = 0.141$ . Figure 7-10 show the range and mean values in response to this question per condition.

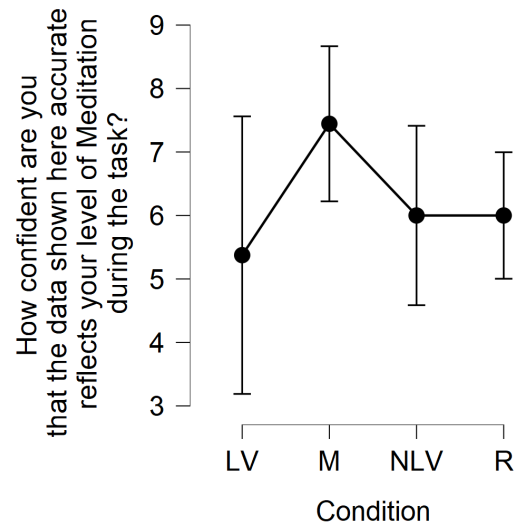


Figure 7-10: Confidence in Accuracy of Data Visualisation - Meditation

**How confident are you that the device [Fitbit] used to collect this data is trustworthy?**

An analysis of variance showed that the effect of the data visualisation on confidence that the Fitbit was trustworthy was not significant,  $F(3, 32) = 1.720$ ,  $p = 0.183$ . Figure 7-11 show the range and mean values in response to this question per condition.

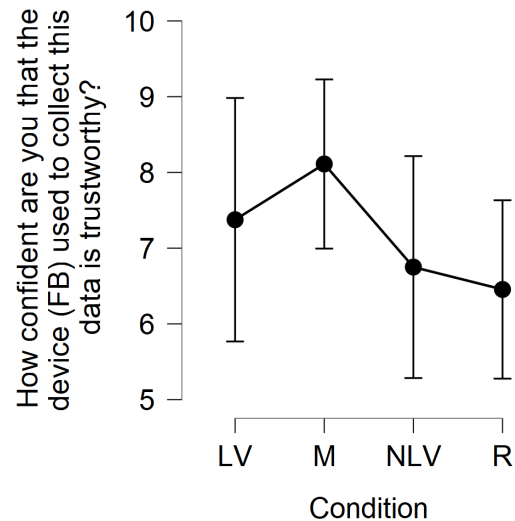


Figure 7-11: How confident are you that the device used to collect this data is trustworthy? - Fitbit

### How confident are you that the device [NeuroSky] used to collect this data is trustworthy?

In relation to confidence that the NeuroSky headset was trustworthy, no effect was seen, similar to confidence in visualisation accuracy as previously seen for Attention and Meditation.

An analysis of variance showed that the effect of the data visualisation on confidence that the NeuroSky was trustworthy was not significant,  $F(3, 32) = 0.482$ ,  $p = 0.697$ . Figure 7-12 shows the range and mean response to this question per condition.

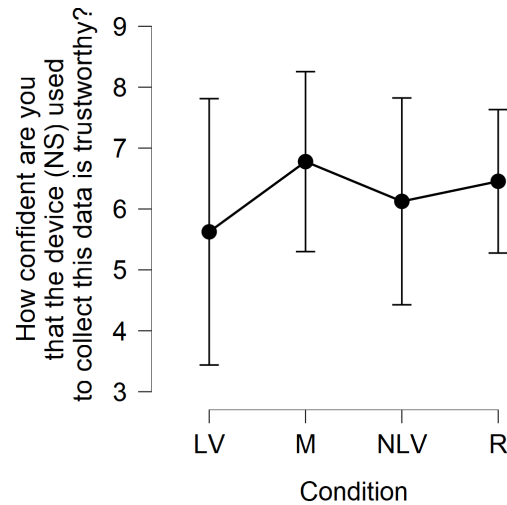


Figure 7-12: How confident are you that the device used to collect this data is trustworthy? - NeuroSky

### 7.3.2 Match Between Visualised Data and Prior Expectations

#### How closely does the data shown here reflect your prior expectation of how your *<metricname>* might change during the task?

Participants in each group were asked to rate how closely the data shown reflected their prior expectation for each recorded metric e.g. *How closely does the data shown here reflect your prior expectation of how your <metric name> might change during the task?* where metric name was either 'heart rate', 'Attention' or 'Meditation'.

For all three of the metrics no significant differences were observed across conditions (see Figure 7-13). The results of an analysis of variance showed that the effect of the data visualisation on how closely the data matched prior expectations was not significant for Attention, Meditation and heart rate:  $F(3,32) = 0.632$ ,  $p = 0.6$ ;  $F(3, 32) = 0.913$ ,

$p = 0.446$ ; and  $F(3, 32) = 1.158$ ,  $p = 0.341$ , respectively.

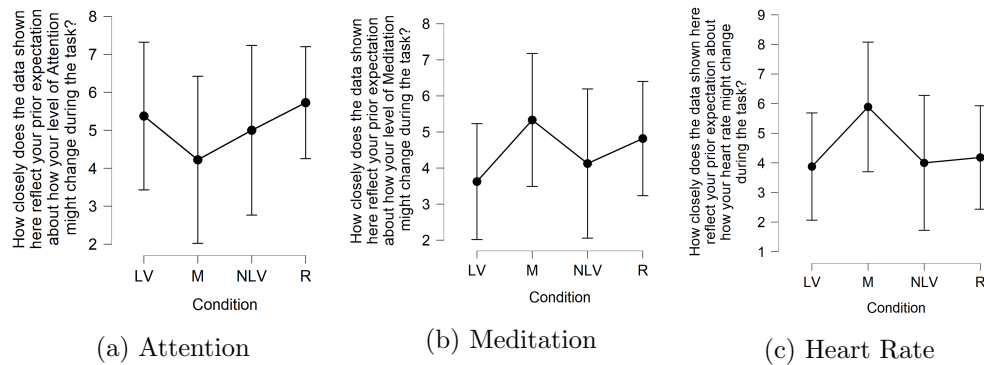


Figure 7-13: Metric Data Visualisation - Closeness to Prior Expectation

### 7.3.3 Exploratory Analysis

Further exploratory analysis was conducted to search for any other factors that may influence trust in the devices being used and values being measured. Based on pre-test responses to a question asking participants to rate their confidence in a device to reflect the level of the metric that is supposed to record we see an increased confidence ratings for Fitbit versus the NeuroSky overall.

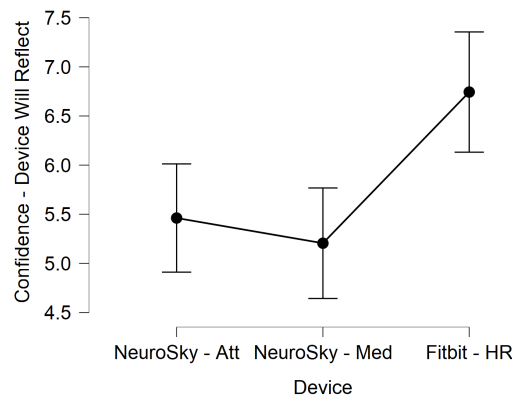


Figure 7-14: How confident are you that this device will accurately reflect your level of... ?

In Table 7.2 we present the results of our correlation analysis with significant correlations ( $p < 0.05$ ). As multiple comparisons are being performed, there is an increased likelihood of seeing a significant result as a result of Type I error. Thus, we present all of those with  $p < 0.05$  but only proffer interpretations for those with p-values is smaller, where  $p \lesssim 0.01$  (highlighted in bold). We do this to limit our interpretation to

those which present the highest significance and thus least likely to be seen by chance due to multiple comparisons.

Variable One	Variable Two	Pearson's r	p
pre_tracking_experience	post_neuro_trust	0.347 *	0.038
<b>pre_fitbit_knowledge</b>	<b>pre_neuro_knowledge</b>	<b>0.583 ***</b>	<b>&lt; .001</b>
pre_fitbit_knowledge	med_delta	-0.354 *	0.034
pre_fitbit_knowledge	post_total_delta	-0.337 *	0.045
<b>pre_fitbit_confidence</b>	<b>pre_neuro_confidence</b>	<b>0.547 ***</b>	<b>&lt; .001</b>
pre_fitbit_confidence	post_fitbit_match_expectation	0.337 *	0.044
pre_fitbit_confidence	post_fitbit_trust	0.418 *	0.011
<b>pre_fitbit_confidence</b>	<b>att_delta</b>	<b>-0.450 **</b>	<b>0.006</b>
<b>pre_fitbit_confidence</b>	<b>post_total_delta</b>	<b>-0.572 ***</b>	<b>&lt; .001</b>
<b>pre_fitbit_confidence</b>	<b>post_fitbit_delta</b>	<b>-0.516 **</b>	<b>0.001</b>
<b>pre_fitbit_confidence</b>	<b>post_neuro_delta</b>	<b>-0.425 **</b>	<b>0.010</b>
<b>pre_neuro_confidence</b>	<b>post_neuro_trust</b>	<b>0.647 ***</b>	<b>&lt; .001</b>
<b>pre_neuro_confidence</b>	<b>med_match_expect</b>	<b>0.672 ***</b>	<b>&lt; .001</b>
<b>pre_neuro_confidence</b>	<b>post_neuro_match_expect</b>	<b>0.471 **</b>	<b>0.004</b>
pre_neuro_confidence	post_fitbit_match_expectation	0.389 *	0.019
<b>pre_neuro_confidence</b>	<b>att_delta</b>	<b>-0.474 **</b>	<b>0.004</b>
<b>pre_neuro_confidence</b>	<b>post_total_delta</b>	<b>-0.500 **</b>	<b>0.002</b>
<b>pre_neuro_confidence</b>	<b>post_neuro_delta</b>	<b>-0.465 **</b>	<b>0.004</b>
<b>total_confidence_pre</b>	<b>med_match_expect</b>	<b>0.595 ***</b>	<b>&lt; .001</b>
total_confidence_pre	post_neuro_match_expect	0.411 *	0.013
total_confidence_pre	post_fitbit_match_expectation	0.416 *	0.012
total_confidence_pre	post_fitbit_trust	0.330 *	0.049
<b>total_confidence_pre</b>	<b>post_neuro_trust</b>	<b>0.567 ***</b>	<b>&lt; .001</b>
<b>total_confidence_pre</b>	<b>post_total_delta</b>	<b>-0.592 ***</b>	<b>&lt; .001</b>
<b>total_confidence_pre</b>	<b>post_fitbit_delta</b>	<b>-0.443 **</b>	<b>0.007</b>
<b>total_confidence_pre</b>	<b>post_neuro_delta</b>	<b>-0.507 **</b>	<b>0.002</b>
<b>med_match_expect</b>	<b>post_neuro_trust</b>	<b>0.723 ***</b>	<b>&lt; .001</b>
<b>med_match_expect</b>	<b>med_delta</b>	<b>0.529 ***</b>	<b>&lt; .001</b>
<b>post_neuro_match_expect</b>	<b>post_neuro_trust</b>	<b>0.561 ***</b>	<b>&lt; .001</b>
<b>post_fitbit_match_expectation</b>	<b>post_fitbit_trust</b>	<b>0.524 **</b>	<b>0.001</b>
post_fitbit_match_expectation	post_neuro_trust	0.344 *	0.040
hr_delta	att_delta	0.377 *	0.023
att_delta	post_fitbit_delta	0.377 *	0.023

\* p < .05, \*\* p < .01, \*\*\* p < .001

Table 7.2: Pearson's Correlations

Participants who self-reported more knowledge of Fitbit also reported more knowledge of NeuroSky ( $r=0.583$ ,  $p<0.001$ ).

Participants who were more confident that Fitbit would accurately reflect the data being recorded also rated their confidence in NeuroSky's ability to accurately reflect the data being recorded more highly ( $r=0.547$ ,  $p<0.001$ ).



Although it does not meet the lower p-value of 0.01, we draw attention to pre-test confidence in Fitbit correlating with higher post-test ratings for Fitbit’s trustworthiness ( $r=0.418$ ,  $p=0.011$ ).

Higher confidence ratings pre-test also correlate with smaller changes (lower directed delta values, positive increase or negative decrease) in ratings between post-test trustworthiness. This correlation is seen for Attention ( $r=-0.450$ ,  $p=0.006$ ) and for changes in total trust ( $p=-0.572$ ,  $r<0.001$ ). Participants who rated their confidence in Fitbit more highly pre-test had a lower change in rating for the Fitbit device being trustworthy post-test ( $r=-0.516$ ,  $p=0.001$ ), as well as the NeuroSky device being trustworthy post-test ( $r=-0.425$ ,  $p=0.010$ ).

These results suggest that a person’s initial confidence in the Fitbit device may be a determining factor rather than the visualisation of the data. The correlation could suggest that initial confidence, rather than whether a visualisation matches expectations is a factor in determining a device’s trustworthiness.

A similar effect was also seen in relation to NeuroSky. Higher participant confidence ratings for NeuroSky correlate with higher trustworthiness ratings post-test ( $r=0.647$ ,  $p<0.001$ ). Higher pre-test confidence in NeuroSky correlated with smaller differences between pre- and post-test trust ratings for Attention ( $r=-0.474$ ,  $p=0.004$ ), NeuroSky ( $r=-0.465$ ,  $p=0.004$ ), and total trust ratings ( $r=-0.5$ ,  $p=0.002$ ).

The recurrence of similar correlations for NeuroSky, as were noted for Fitbit, suggests that that the same factors may hold true across different sensing devices, rather than being specific to only Fitbit.

Participants who rated their confidence in NeuroSky more highly pre-test, rated their confidence in the Meditation visualisation matching their expectation more highly post-test ( $r=0.672$ ,  $p<0.001$ ). Participants also rated the combined value of NeuroSky Attention and Meditation matching their expectation higher post-test when they rated their confidence in NeuroSky higher ( $r=0.471$ ,  $p=0.004$ ). These results suggest that higher levels of confidence in a device may result in people being more inclined to trust the values being shown as accurate. However, given that we only see this for a single metric and not for any of the others, there may be other factors to take into consideration.

Higher total confidence pre-test was found to correlate with increased confidence that the visualisation of Meditation accurately reflected the level of Meditation during the Stroop test ( $r=0.595$ ,  $p<0.001$ ). Total confidence pre-test also correlated with higher

post-test NeuroSky trust ( $r=-0.567$ ,  $p<0.001$ ). Higher total confidence pre-test correlated with lower differences in pre- to post-test total trust values ( $r=-0.592$ ,  $p<0.001$ ), Fitbit trust values ( $r=-.443$ ,  $p=0.007$ ) and NeuroSky trust values ( $r=-0.507$ ,  $p=0.002$ ). Higher ratings for the Meditation visualisation matching the participant's expectations correlated with increased changes in Meditation ratings values ( $r=0.0529$ ,  $p<0.001$ ).

We also note that when condition is not taken into account, participants who self-reported that the visualisations more closely matched their expectations rated their post-test trust in the devices more highly both for NeuroSky ( $r=.561$ ,  $p<0.001$ ) and Fitbit ( $r=.524$ ,  $p=0.001$ ). It is not clear which variable is the influencing/causal factor, whether increased trust is more likely to make someone believe that the data matches their expectations, or whether it is the data matching self-reported expectations that engenders increased trust.

### 7.3.4 Open-ended Response Analysis

Participants were asked '*Why?*' for each of the following questions:

#### Pre-Test

- How confident are you that this device will reflect your heart rate?
- How confident are you that this device will reflect your level of Attention?
- How confident are you that this device will reflect your level of Meditation?

#### Post-Test

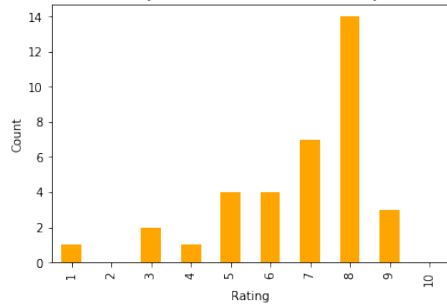
- How confident are you that the device [Fitbit] used to collect this data is trustworthy?
- How confident are you that the device [NeuroSky] used to collect this data is trustworthy?

Participant's open-ended responses were inductively coded and analysed. Initial codes were generated on the first pass of participant responses and the following similar question used these codes and were expanded as necessary.

#### How confident are you that this device will reflect your heart rate?

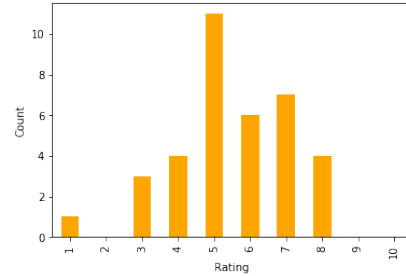
Codes related to comments made in relation to Fitbit's ability to reflect heart rate can be seen in the codes presented in Table 7.3.

How confident are you that this device will reflect your heart rate?



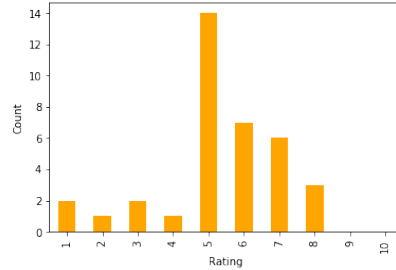
(a) ...heart rate? - Fitbit

How confident are you that this device will reflect your level of Attention?



(b) ...level of Attention? - NeuroSky

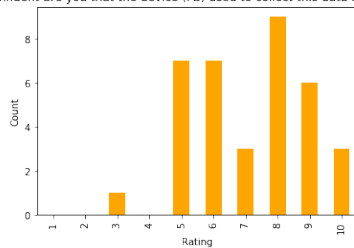
How confident are you that this device will reflect your level of Meditation?



(c) ...level of Meditation? - NeuroSky

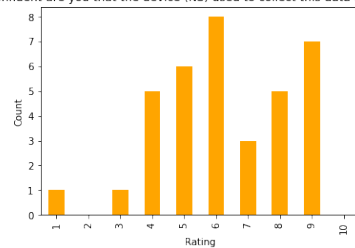
Figure 7-15: How confident are you that this device will reflect your...?

How confident are you that the device (FB) used to collect this data is trustworthy?



(a) Fitbit

How confident are you that the device (NS) used to collect this data is trustworthy?



(b) NeuroSky

Figure 7-16: How confident are you that the device used to collect this data is trustworthy?

Code	Count
Other people are using it	6
Errors possible	4
No experience	4
Tested Device Self	3
Belief/Trust in the tech	3
Unsure of device/claims	3
Previous experience (Negative)	3
Advertised as such	2
Science (Negative)	2
Science (Positive)	2
Easy to measure	2
As accurate as other technologies	1
Device has been around a while	1
Prior Awareness	1
Don't Know	2
No Answer	2

Table 7.3: How confident are you that this device will reflect your heart rate - why?

Based on these comments it appears that one of the main considerations when determining confidence is whether other people, either in terms of the general population or by someone known personally, are using the device. Participants expressed this is a variety of ways, such as P25 who stated that:

*Because many people buy [F]itbits, so I assume there was a lot of research done into the efficacy of the tracking, as some buyers would have checked the accuracy and left reviews. (P25)*

This type of comment was similarly echoed by P24 who said that as “*The devices are very popular and wouldn't be so if they were not accurate*” (P24). P23 said they “*guess it works*” (P23) because “*a lot of people are use it*” (P23). However, not all instances were based on large numbers of people using the device and instances were more based on knowledge of a specific person, one participant based their assertion on basis that their “*wife uses one or similar*” (P27) and P37 said that they “*hadn't heard anyone saying that these things fundamentally don't work*” (P37). Another participant suggested that their confidence was influenced by a general awareness of such technologies, “*I'm aware of Fitbit style devices that can monitor heart rate and feed to a mobile device*” (P21) rather than specifically being aware of others using the device.

These responses suggest that acceptance of a product into the mainstream may play a

role in influencing whether people will trust that a device does what it as advertised to do, and will do it with sufficient accuracy. However, there is likely some interplay between these two things. A device may be more likely to be accepted into the mainstream based on a number of factors and its perceived ability to accurately record and reflect accurate values may be one of the factors that increases adoption to the mainstream in the first place.

Another factor that influenced participants' confidence was having had the ability to test the device, or a similar technology, themselves previously. These participants compared the values they were getting to what they perceived to be a reliable source and have used that knowledge to make a decision about how confident they are in the device's abilities. P8 mentions previous experience with a Fitbit, *"I use a fit bit regularly and the heart rate matches up with the activities I am doing"* (P8). P19, however, makes a decision based on experience with a different technology:

*I am also using a smart watch for measuring my heart rate and else. When I compared the result with the treadmill, the result is similar. So I think it would be the same for Fitbit Versa as well.* (P19)

P6 made a similar comparison to a previous experience with another technologies but was slightly less confident (6 vs. 9) than P19, saying that *"the measurements were different within one period, and not always accurate"* (P6). These comments appear to suggest that having the ability to test the device and validate its accuracy may play a role in confidence. However, it is interesting to note that P19 appears to draw their conclusion from experience with a a similar, but different, sensing technology. This suggests that there may be some transference of confidence between technologies such that one might assume that if one piece of technology can adequately sense data then other technologies that sense the same data should be of similar ability.

Participants who had no previous experience clearly stated this as such in their comments. However, previous personal experience with a Fitbit, or similar technologies, allowed some participants to use that experience in their assessment of the devices they were using. Although often this was presented negatively, for example having a previous experience in which *"other devices did not work well"* (P26) and *"Inaccuracies due to previous usage...I've been in spin classes working v hard yet my hr is only showing as 110bpm, when i can feel it is more like 160+"* (P28). This is somewhat related to the previous responses but rather than relying on the experience of others, people are able to use their own personal experiences to make inferences about a sensors reliability. It is interesting to note that those with previous experience are qualitatively less positive

in their responses.

Rather than basing their confidence on other people using the technology some participants' comments suggest that there is an innate trust or belief in the technologies that are available. Some participants were less descriptive in their responses such as P7 who just said *"Because I believe the technology"* (P7). Whilst some based their belief on the fact they felt the description sounded plausible *"The description of how it works seems detailed enough to let me think the producers know what they are doing"* (P32). These comments suggest that device manufacturers also play a role in convincing consumers that the validity of their claims is believable. Therefore, how a device is described and whether that description is believable is another factor that may influence confidence in the device.

A cause of doubt for some participants was the error in accurately capturing the data. Participants explained being aware that there was room for error when sensing heart rates. Two participants just highlighted that there was always a possibility of error when measuring signals generally. Two participants specifically referred to sensor placement as a factor that would influence data errors. P18 said that *"there is more error associated with wrist-based monitoring versus chest-based"* (P18). P38 rather than comparing different modes of wearing a sensor referenced sensor placement on the wrist and sensor quality as factor that influenced their confidence, *"Depends on position on wrist and sensors quality - i don't not [sic] expect a very accurate measurement"* (P38).

These concerns, however, require some prior knowledge in how signals are acquired and understanding of how sensor placement could effect signal quality and thus the accuracy of the data captured. It is not clear from the participants comments whether this knowledge was acquired from having experience with a variety of sensing technologies or whether these are based on knowledge that we may not consider an average user to have.

Other participants believed that the science sounded credible ( *"Photoplethysmography sounds like a technique that would work, as I have heard of similar things in my Biology degree"* (P24)), however, not all participants thought of the science in a positive light but rather stated that they felt the technology could be flawed because of the science used, *"Different skin colors might behave differently. Some light might be absorbed by other biological matter in the wrist. Some may even dissipate in the environment"* (P5). This type of comment suggests that people with specific knowledge may take specific factors into consideration when assessing their confidence in a device, rather than blindly accepting what a manufacturer has stated.

### How confident are you that this device will reflect your level of Attention?

Code	Count
Hard to Measure	7
Science (Negative)	5
Not enough information	5
Belief/Trust in the tech	4
No experience	4
Advertised as such	3
Not accurate	2
No confidence	1
Errors possible	1
Previous Experience	1
Previous exp. (Negative)	1
Aware of such technologies	1
Aware of more complex systems	1
Setup	1
Need to test	1
No Answer	2
Don't Know	1

Table 7.4: How confident are you that this device will reflect your level of Attention - why?

Given the lower ratings for both of NeuroSky’s metrics the comments reflect slightly more negatively than Fitbit’s comments. These comments (see Table 7.4) therefore include believing that Attention would be hard to measure, “*I am aware that many EEG signals can be inconsistent or reflective of things other than just attention for example*” (P25), or that the science behind it may be inadequate, “*It perhaps oversimplifies complex concepts into rather arbitrarily variables with no explanation of the algorithms used to calculate these*” (P41). These comments are not quite the same as seen in relation to Fitbit in that they comments are associated with the difficulties that come with measure EEG and providing a value based on that. Nobody particularly stated that heart rate was difficult to measure. This is possibly a valid reason in that attempting to resolve EEG data into a singular value is not a simple problem to solve. However, it might be argued that the way in which Fitbit attempts to use light to detect blood flow is not an easy problem either.

Participants also expressed a sense of the devices being inaccurate, stating “*I don’t think it’s accurate enough, and I think people can judge their level of attention by themselves*” (P34). Other comments suggest that due to the way in which the sensor

works may be a reason for the inaccuracy, *“Detector still lies outside of the scalp, measurement might not be entirely accurate”* (P17). These comments suggest that there is an unease with the accuracy they perceived the device being capable of achieving. With the exception of mentioning sensor location it is not entirely clear how participants came to this conclusion. P23, however, seemed to think that the device appeared too simplistic to be accurate saying, *“The ones that I’ve seen that measure similar variables seem to be more complex”* (P23).

Participants said that they did not have enough information, based on the description and lack of previous experience, to rate their confidence more highly:

*I do not have enough information about inner processes of a person used to measure the level of calmness and relaxation. I would like to know exactly what will be measured (pulse, waves, etc, etc).* (P10)

This suggests that some people will question or openly accept that they do not have sufficient information to make a decision about a device. However, despite this some of the same types of comments that were associated with Fitbit were mentioned in relation to NeuroSky’s ability to measure their level of Attention including having an innate trust or belief in the technology. For example, it was suggested that despite lacking familiarity with the technology they were still willing to believe what was being described:

*I have no reason to not believe the description above, but I am not familiar with the technology at all and have not come across these concepts of measuring Attention/Meditation before.* (P11)

Similar comments suggested that they had no reason not the trust that the device did what was advertised, *“The text above makes the claim and I’ve no reason to doubt it”* (P37).

### **How confident are you that this device will reflect your level of Meditation?**

As with Attention, comments related to NeuroSky (see Table 7.5) being able to reflect their level of Meditation, participants’ suggested reasons were largely negative and included not having enough information to rate their confidence more highly, feeling that that specific metric would be hard to measure or that the measurement would include error or problems in terms of accuracy. Participants commented on their lack of information saying, *“I’m not sure what EEG spectrum represent and how this translates into level of meditation. It also seems a more vague concept than attention”* (P8) or



Code	Count
Not enough information	10
Hard to Measure	7
Errors possible	3
Not accurate	3
Science (Negative)	2
Belief/Trust in the tech	2
Unsure of device/claims	2
Advertised as such	1
No confidence	1
Previous exp. (Negative)	1
Aware of such technologies	1
Need to test	1
No Answer	2
Don't Know	2

Table 7.5: How confident are you that this device will reflect your level of Meditation - why?

simply, *“i don’t know how it exactly works”* (P15). This again reiterates the idea that not all people are open to being highly confident in a technology. One participant said that the quantification might be difficult to achieve but that in combination with other devices they felt it might be easier, *“Combined with other aspects like blood pressure, heart rate, level of oxygen in someone’s body it might be relatively easier to measure the calmness. There again is the problem of quantifying and actually measuring it”* (P5). P13 was not completely confident in the devices abilities stating *“I believe it will work to reflect the level of meditation, but I believe there will be noise and error exist”* (P13).

Participants also suggested that they had previous negative experiences with this type of technology in the past with P14 reiterating the comment they had made in relation to the ability of the device to measure Attention, *“I had previous experience with this device and I didn’t always agree with its measurements and how I thought I felt”* (P14). This comment additionally highlights the inconsistencies between what a device may output and how the user might differ in opinion, particularly when the things being measured may be subjective.

The less cautious comments by participants suggested having an innate belief in the technology and having no reason to doubt that the technology works as advertised. One participant felt that *“[t]he device seems to measure variables that could indicate the level of meditation”* (P18), whilst another stated *“[t]he device seems reliable”* (P19).

These comments were very much in the same vein as those for Attention highlighting a possible assumption about the reliability device based on preconceptions, or limited information from a description, that people may be making.

**How confident are you that the device [Fitbit] used to collect this data is trustworthy?**

Code	Total	LV	NLV	R	M
Output matches expectation	13	3	2	4	4
Output didn't match expectation	10	2	3	4	1
Belief/Trust in the tech	2	1	1	0	0
Rough Indication	2	1	1	0	0
Easy to measure	2	1	1	0	0
Errors possible	1	0	1	0	0
Unsure of device/claims	1	0	0	1	0
Trust despite not meeting expectation	1	0	0	1	0
No experience	1	0	0	1	0
Other people are using it	1	0	0	1	0
Device has been around a while	1	0	0	0	1
Well-known brand	1	0	0	0	1
Output plausible	1	0	0	0	1
Don't Know	2	0	1	0	1
No Answer	1	1	0	0	0

Table 7.6: How confident are you that the device [Fitbit] used to collect this data is trustworthy - why?

In relation to the trustworthiness of Fitbit as a device most comments were related to the data visualisation either matching or not matching their expectation (see Table 7.6). Of the participants who made comments in this regard most (13) believed that the output did match their expectation, whilst 10 participants felt it did not meet their expectations. One participant specifically referred to the output being ‘plausible’ rather than overtly stating it matched or did not match their expectations. Some participants expressed very clearly that the heart rate matched what they expected, *“as the reflection of the heart rate is similar as I expected”* (P7, NLV) whilst others were more varied in their response, *“I would expect more variations and upward trend in the heart rate in fifth section. In other sections it more or less meets my expectation. Because of these reasons I have somewhat above average trust in the device”* (P5, LV). For those whose data did not match their expectation there was a similar variety of expectations with some participants providing specific details about where the data did not match, *“I feel quite nervous during the last trial, I think it will rise in the last trail. But it’s*

*quite flat like the first two trail*” (P13, NLV) and others being less specific, *“i think my heart rate was fluctuating more than shown by the graph”* (P16, LV). Other participants reflected to their own experience of how their heart rate should be and how that did not appear to be reflected in the data, *“My heart rate is normally at around 60-65 so I am surprised it would reach 90 - also surprising as I would expect it to increase as the task for harder - or peak for the beeps”* (P25, R).

Participants again related their trust in the device to innate trusts or belief in the technology, as well as believing that heart rate should be easy to measure. A couple of participants felt that the device was trustworthy enough that it could give a rough indication of heart rate.

**How confident are you that the device [NeuroSky] used to collect this data is trustworthy?**

Code	Total	LV	NLV	R	M
Output didn't match expectation	15	4	3	4	4
Belief/Trust in the tech	7	2	2	2	1
Output matches expectation	3	0	1	1	1
No experience	3	0	1	1	1
Subjective	2	1	1	0	0
Hard to Measure	2	1	0	1	0
Unsure of device/claims	1	1	0	0	0
Errors possible	1	0	1	0	0
Disconnect with metric name	1	0	1	0	0
Trust despite not meeting expectation	1	0	0	1	0
Looks trustworthy because of variability	1	0	0	1	0
Output plausible	1	0	0	0	1
Ground Truth - Not Aware	1	0	0	0	1

Table 7.7: How confident are you that the device [NeuroSky] used to collect this data is trustworthy - why?

For NeuroSky a large portion of participants’ (see Table 7.7) felt that the data visualisation did not meet their prior expectation. Participants felt that the data was too noisy, *“It feels like the data is very noisy, meditation output does not match my expectation”* (P3, M) or that they believed that the data should be more constant *“I thought I would be in constant focus”* (P14, R). One participant suggest that the variability of the data made it very difficult to interpret the data as it was *“hard to relate it back to specifics within my experience of the task - and does not follow the trend I expected”* (P25, R). However, another participant suggested that despite the

noise there appeared to be some visible trends in the data, *“Whilst the graph does show something of a trend downwards, it seems to vary wildly within that range. I certainly didn’t feel like my calmness varied by that much, and especially not with that frequency.”* (P21, NLV). However, not all participants specifically mentioned that the data was too noisy, or unclear in its presentation but rather expressed more generally that they expected *“higher attention and meditation in first few sections. As the speed of the task increases...”* (P5, LV).

These comments appear to show various types of situations which may cause people to be less trusting in sensing technologies. Mainly this appears to be related to data not matching people’s expectations of how the data should look.

Despite these previous comments, a few participants felt that the data visualisation met their expectations saying the graph *“shows what i felt and similar with my expectation”* (P30, M).

Other comments also followed the previously seen tendency for participants to trust, or believe, that the technology is correct. Participants expressed these feeling as though the device *“seems more reliable to me, and I somehow really felt the same way when I took the test”* (P19, M) or simply, *“i trust the graph of the finding”* (P16, LV). This expression also appeared to be presented as due to not being familiar with the technology, they *“just have to trust it”* (P33, R).

## 7.4 Discussion

### 7.4.1 There will be a significant increase in user trust in data when the data more closely reflects users’ prior predictions about data values.

Our initial hypothesis was that there would be a significant increase in user trust in data when the data presented more closely reflects prior predictions about data values. It was seen that participants in Condition M showed an increased level of confidence that the values accurately reflected their heart rate during the task. This difference was only noted between NLV-R and M-R, the effect of rounding is discussed in the following section.

No significant difference was seen between groups in respect of Attention values. This may be due to the lack of experience of how these values should look resulting in no significant differences being noticed. Another possibility is that that despite the

manipulation the visualised data was still too noisy for the effect of the manipulation to truly be sufficient.

Participants in Condition M rated the closeness of the values visualised, for all metrics, in similar ranges as those in other conditions. This may be interpreted in two ways. The first is that the effect was sufficiently unnoticeable that it was not detected by participants in group M, despite seeing a noticeable effect in the ratings for confidence in the visualisation of heart rate data. Alternatively, that the effect of manipulation was not sufficiently noticeable that participants were not more likely to rate the values as being closer.

Qualitative responses for Fitbit and NeuroSky suggests that output matching expectations is a factor but it may be that unfamiliarity with the data could be an underlying issue in this respect. For Fitbit that was an almost even mix of participants who said that their data visualisation did or did not match their expectation. However, for NeuroSky none of the participants specifically mentioned the data matching their expectations and 15 participants stated that their data did not match their expectations.

Rather than people questioning data that does not conform to their prior expectations we find that participant's expectation of the data played a limited role in determining whether a device was trustworthy. Although based on different data sources, this is somewhat different to previous research that suggested that people tend to trust data more that conforms to their prior expectation (Y. S. Kim et al., 2018).

#### **7.4.2 There will be a significant difference in user trust in a personal informatics device and the data it produces according to the precision at which it is presented within a visualisation. Therefore rounded values will be less trusted than precise values.**

With respect to heart rate there was a difference noted between participants in conditions NLV-R and R-M. Ratings in respect of confidence of the data visualisations being an accurate reflection of the participant's data were lowest in group R. Although this one result agrees with our hypothesis, there may be other factors at play meaning the results were not seen across metrics. For example, the noisy nature of the EEG data may have made the rounding of values less clear than it may be for heart rate which contains less variations making the heart rate values appear too discrete rather than continuous values as one might expect from a heart rate value. This lack of difference is in line with previous research in relation to the effect of data visualisation precision on trust (Olsen, 2018).

#### **7.4.3 Trust will differ according to whether the data is presented synchronously (in real-time) vs. asynchronously (post-task), based on participants' ability to check 'in the moment' whether the data they are seeing corroborates their expectations.**

We noted no significant difference when participants were able to see a live view of their data being recorded (Condition LV) and those who were not able to see their data being recorded during the Stroop test (Condition NLV). This lack of significant difference was noted for all three of the recorded metrics.

#### **7.4.4 Trust in the sensor data will differ based on the participants level of knowledge they have about how a particular sensing technology is measuring and capturing data.**

We found no significant correlation between people's self-reported level of knowledge of a device and their confidence in the data visualised or trust in the devices. This may be due to a lack of diversity in knowledge ranges in the group of participants, further analysis with a larger variety of knowledge levels might produce a different outcome.

#### **7.4.5 Trust is Influenced by Initial Confidence**

Our further analysis of participants' responses seems to suggest that the basis of confidence and trust may be set prior to people reviewing data from a sensing technology. Trust in a device may not be influenced specifically by how the data is presented but rather by a person's initial levels of confidence in the device they are using. This result when considered in relation to previous work by Hollis et al. (2018) and Muir (1987) perhaps highlights a deference to technology. Whereby peoples' initial confidence in a device is not influenced by the data presented but rather is a pre-decided factor that influences their willingness to trust the technology.

When we consider that trust is linked to a person's prior confidence in the device itself, our qualitative results suggest that this confidence is built upon a trust that other people are using the technologies and thus are considered by others to be reliable. In addition to this our qualitative results suggest that inherent trust or belief in technology also plays a role in determining people's initial confidence. Reasons for not trusting technologies appear to be related to either previous experiences with the technologies, previous situations having used and found the technology not to be reliable, and having prior knowledge about the issues that might be encountered in recording physiological signals.

In relation to NeuroSky that majority of reasons to not be confident in the technology was due to a lack of information. Interestingly no participants mentioned other people using this technology, which is likely to be expected as it is not a widely-known device but other similar reasons to trust in the device were seen, such as an innate belief in technology.

Previous work highlights the features that people may base their expectations of emerging technologies as a framework of user expectations. Drawing on work from multiple disciplines and user feedback, Olsson (2014) posits that with a lack of information about service promises from a manufacturer, people may rely on past experience, word-of-mouth, personal needs and social factors, including others' opinions and experiences.

As has been seen previously, people show no significant ability to determine when a sensing technology is not presenting true values (Huppert et al., 2019). Additionally, it has been seen that misunderstanding how a technology senses the presented values can lead to an eroding of trust in technologies (Yang et al., 2015). Although work aims at increasing user trust in devices such as by being more transparent to users about how specific values have been achieved (Jaimes et al., 2013) and whether this increases user trust. There are still unanswered questions about what initial confidence and trust in the device itself plays. Given that people may have initial levels of trust based on confidence, rather than ceaselessly increasing user trust, it may be beneficial to provide a means confirm or realign confidence and trust; enabling people to form a truer understanding of how and why that data looks the way it does, and look to provide opportunities to enhance people's skills to correctly investigate the data such that they will be aware of devices that are not providing accurate levels of information so that they can make informed decisions about what data to trust.

## 7.5 Limitations

The majority of participants in this study were recruited from within a university setting, as such the results may not be reflective of the broader population of people who may wish to collect, record and analyse data about themselves. Our initial intention for this study was to recruit 80 participants to ensure there was significant power behind the presented results. Having not reached that number the results should be treated with

some caution, although they provide an interesting starting point for future researchers.

## 7.6 Conclusion

In this chapter we described a study which sought to understand whether trust in a self-tracking technology might be influenced by how that device's data is visualised. In particular we hypothesised that this trust might be increased if the data more closely matched people's prior expectations of how the data would change. As well as whether trust might be influenced by being able to view the data recording in real-time, whether prior knowledge was a factor or whether precision might play a role. We found significant differences for Fitbit and heart rate but not for the NeuroSky device. The significant difference for Fitbit and heart rate values was only significant in relation to data that was presented with less precision.

Rather than these appearing to be particularly significant factors in determining trust we now suggest that initial confidence in a technology may be a determining factor that influences whether a person will trust the technology and accept that their data being visualised as being accurate. These initial determining factors in relation to existent technologies appear to be based around societal trust in the technologies. Whilst with emerging technologies the trust appears to be based around deference to technology itself, although qualitatively people appear to state they lack the information to make an informed decision, and knowledge that the values may be hard to measure.



## Chapter 8

## Conclusion

## 8.1 Thesis Summary

The research conducted as part of this thesis aimed to better understand the challenges that may arise from the increasing availability of sensor data and its potential use in personal informatics tools, or by users who wish to use the data they are capturing to better understand themselves, or aspects of their lives. We started by conducting an exploratory study to better understand the ways in which users may use our chosen technology probe, a NeuroSky MindWave, as a basis to inform the development of a multifaceted PI tool that was used in later studies. This exploratory study uncovered a set of initial challenges related to how people attempt to make sense of EEG data in the context of a novel sensing technologies and PI.

We further explored these challenges through two in-the-wild studies. We developed a multifaceted PI application that included features and design changes built upon the results of each previous study. The first in-the-wild study was open to all healthy participants and provided insights into how these tools would be used by general users. For the second in-the-wild study we recruited participants living with ME/CFS, CF, and ICF. This contributed a further understanding of how people with a specific life-aspect to track may attempt to use and make sense of novel PI systems. This second study also tested challenges and benefits of using two features designed to aide in data analysis. These features provided the user with the ability to review data using automatically generated statistical insights from recorded data, presented in the style of a news feed, or via a set of predefined statistical queries, presented as a Query Area. Our final study sought to better understand the impact of visualised data on users' trust in the data produced by a sensing technology, based on a thread related to trust that was discussed throughout our studies.

## 8.2 Contributions and Future Research Directions

In this thesis we presented finding over the course of five chapters, the contributions of which can be described broadly as design considerations and challenges for designing PI tools which use data from emerging sensing technologies. Additionally, we presented a contribution about the formation of trust in sensing technologies used within PI.

### 8.2.1 Four Challenges of Cognitive Personal Informatics Systems

In Chapter 3, we presented findings related to the use and design of applications to support 'Cognitive Personal Informatics' (CPI). These challenges were:

- Challenge 1: Addressing the indirection between meaningful psychological states and brain activity data
- Challenge 2: Supporting diverse tracking styles
- Challenge 3: Encouraging exploratory & enquiring approaches
- Challenge 4: Overcoming misconceptions & lack of understanding about the brain

These challenges contribute a better understanding of the areas that require consideration when developing CPI systems. *Challenge 1* highlights an apparent gap between users' expectations of a CPI systems and current capabilities, such as limitations in detecting specific mental states using consumer BCI devices. We found users did not tend to be concerned with the accuracy of the system or the scientific basis on which the data was formed, in line with prior work by Lawson et al. (2015). Thus we suggest that there is an onus of the developers of novel sensing technologies, and respective PI systems, to ensure that users are not misinformed by the system and adequate information is provided. For *Challenge 2*, we suggest that CPI systems should support diverse tracking styles, in line with prior research's views of how people track data using PI systems (Rooksby et al., 2014). Although one current limitation to this is providing people with a directive tracking style with appropriate values to 'target', due to a lack of appropriate guidelines in relation to cognitive states, such as there is for physical activity (e.g. 10,000 steps per day). Further research is required in this area to determine appropriate values and the ability for these to be determined through EEG recordings. Therefore, in *Challenge 3*, we highlight the importance of not suggesting that the answer to a medical problem can be found at the press of a button and agree with prior work that highlighted the importance of avoiding deterministic diagnoses but rather advocated for exploration in self-tracking (Ayobi et al., 2017). Finally in *Challenge 4*, we highlight the need to further educate and inform users with 'brain knowledge', suggesting that developers should consider the diverse ranges of knowledge that users of their systems may have and the potential for misheld beliefs to interfere with a user's ability to interpret their recorded data.

These challenges were used as the basis for guiding our further research, the contributions of which are described in the following sections. These contributions can be described broadly as challenges and design considerations for PI systems, and challenges related to sensing technology accuracy and trust.

### 8.2.2 Challenges and Design Considerations for the Design & Use of PI Systems

In Chapters 4, 5 and 6, by providing participants with novel sensing technologies and a full-featured PI data collection and analysis tool we were able to contribute design considerations for designers, developers and researchers of PI tools.

We found that participants were keen to use sensing technologies that were discreet, due to concerns about what others may think. Additionally, we find that users were conscious that tracking tools should not prevent them from engaging in their day-to-day lives. Prior work suggested that aesthetic is not a factor in the adoption of wearable technologies Chuah et al. (2016), however our results suggest this may not be the case universally. It also suggested that our initial assertion that novel sensing technologies in the future will be unobtrusive (Chapter 2, Section 2.3.4) holds true - as consumer demand will likely make this so. This result may present a limitation in our choice of technology probe, however given the desire to provide participants with a sensing technology that was not yet mainstream, this was a necessity. Our choice of sensor meets several other characteristics we defined (Chapter 2, Section 2.5).

We found that participants were eager to engage with novel sensing technologies, but felt they may be lacking in prerequisite knowledge to analyse their data. Ding et al. (2021) similarly found that it is no longer the case that users of PI systems are guaranteed to have requisite knowledge to understand their data. This further echoes our initial challenges of CPI systems, where we suggest that users may lack knowledge and drives home that users also feel the need for knowledge when attempting to make sense of their data within the context of PI. Previous research has seen the inclusion of domain experts in the development of PI systems (Epstein et al., 2020), however we suggest that developers or designers could engage with knowledge transfer from domain experts to end users and not just use their input in the design of PI systems. For example, there could be more in-depth information sections included in systems, domain-related walk-through tutorials, or ‘hints’ that explain what specific data means or how to interpret data visualisations.

In our initial in-the-wild study we found that participants showed preferences for reducing the amount of analysis they had to do, or placing less onus on the user to extract information from their data (Chapter 4). This is in line with previous research suggesting that users may suffer from cognitive overload when presented with too much information in PI systems (Katz et al., 2018). As such we suggested that PI systems should find ways to simplify insight extraction. We explored this further in our sub-

sequent study (Chapter 6) to understand whether automated analyses of recorded data was beneficial and what challenges may be faced by users of systems with this style of insight extraction. Based on the results of this latter study we contribute four design considerations when developing PI tools that aimed to reduce the burden of data analysis: *Designing for Exploration and Enquiry*, *Designing to Uncover Insights - Volume & Variety*, *Designing for the Sceptic*, and *Designing For Context-based Analysis*.

We suggest that when *Designing for Exploration and Enquiry*, both a ‘News Feed’-style interface and Query Area of predefined analyses provides user with opportunities to further investigate their data, outside of more traditional summary overviews of data. We suggest Query Areas are likely more suited to those with prior questions that they are seeking answers to. The Insight Feed on the other hand provided user with no specific questions to answer a means of bringing interesting information to the fore. Users with no initial questions were able to use the insights as a starting point for investing their data deeper using the Query Area.

One negative impact of providing insights was that there were too many provided, leading participants to be unaware that certain insights were even being generated. Prior work also highlights issues related to information overload in PI (Bernstein et al., 2010). When *Designing to Uncover Insights - Volume & Variety* we suggest that there is a balance to be achieved between providing all of the available information or only providing subsets of information. Alternatives such as reducing the frequency at which insights are generated or increasing the level at which an insight is deemed ‘interesting’ (statistical significance). We highlight that despite providing our participants with a means of rating generated insights positively or negatively, this feature was not interacted with much. Similar issues related to low rating interactions have also been highlighted in relation to other services, suggesting this may stem from users lacking information about how ratings influence recommendations (Alvarado et al., 2020). We therefore highlight that this may not be an appropriate way to gain feedback that could be used to filter insights based on recommendations. Further research could seek to determine if there are alternative means for building recommender systems that can work within the context of uncovering interesting insights from PI data.

We found in our study that users tended to question generated insights when they did not match their expectations. This adds another consideration to prior work which found that users were frustrated by the inclusion of ‘obvious’ insights (S. Jones & Kelly, 2016). We suggest when *Designing for the Sceptic*, designers should consider whether there are ways to reduce scepticism. However, further research may be required in this area to provide concrete solutions to the problem, this problem may also be linked to

scepticism in technology itself (H. C. Kim, 2015).

Our design consideration, *Designing For Context-based Analysis*, suggests that users tracking for self-understanding wish to be able to analyse data contextually and not in a vacuum. This consideration aligns with prior work, which suggests the lack of context can lead to difficulties uncovering insights (Choe et al., 2014) and contextual information providing necessary anchors on which analyses can be framed (Raj et al., 2019). We suggest therefore that analyses based on contextual factors will likely provide benefit to users. For example, users might be provided with a means of selecting and analysing data by structuring queries in the form of, 'Show me my mental effort on the days where I go to the coffee shop'. Alternatively, and in combination with automatically generated insights, it may be beneficial to sub-divide insights generation to include contextual factors automatically, rather than attempting to analyse just the values alone.

Further to these general design consideration we provide further suggestion, in particular for designing PI tools for those living with chronic conditions, who may undertake tracking in a more diagnostic or directive style of tracking than those tracking in a documentary-style of tracking. For users who wish to track because of a particular condition, we highlight the importance of designing tools that take into consideration particular effects or symptoms that may make sense-making difficult for the user. Prior work had already highlighted this (Davies et al., 2019), therefore we concur with this work and highlight the importance of engaging with participants in the design of tools to ensure that the needs of the target user are met and do not increase or exacerbate symptoms of their condition.

Additionally, for the users living with ME/CFS, CF, and IF that participated in our study there was a desire to be able to support tracking data over the longer-term. We argued that this is not well-supported with many commercially available sensing technologies. Although some research suggests that keeping users sufficiently engaged for long-term tracking may be problematic (Rapp & Cena, 2014) and this is a valid concern that requires further research, we suggest that certain sets of users may be inclined and willing to track longer-term of their own volition. Therefore, developers and designers of PI tools should ensure that their platform supports longer-term visualisations and appropriate access to longer-term data.

Another factor that participants struggled with was knowing how the data related to them, having no previously recorded data, and how their data compared to other or set appropriate targets. Häkkinen, Colley, Inget, Alhonsuo, and Rantakari (2015) has

suggested that available self-tracking tools lack sufficient features for target setting, so there is still room for further exploration in relation to user interfaces. However, more directly related to the data itself, Feustel et al. (2018) has suggested that it may be possible to use cohort data to make comparisons against. Therefore future work may consider the possibilities of sourcing cohort data for the population that the tools is aimed at helping. Considering the expectation that the amount and availability of self-tracking data is likely to increase in the future, it may be that a) people will have prior data to compare themselves against or b) there may be sufficient other people tracking data that we are capable of finding surrogate data for an individual lacking their own.

We also suggest that the design of PI systems should provide a means of interrogating collected data that considers the variety of people and purposes that the system may be used for. This includes people with varying levels of knowledge, both in terms of their ability to interpret data visualisations and statistical information, and technical knowledge. Based on our initial consideration of the increasing availability and use of emerging sensing technologies for PI that this consideration is important because self-tracking may no longer be used by only those with the prerequisite knowledge to undertake these analyses, but simplification of designs may lead to those wanting to dig deeper into their data to be unable to do so. Supporting these considerations may also include alternative means of interacting with PI systems, examples such as voice control (Young-Ho Kim et al., 2021) and virtual reality (Millais, Jones, & Kelly, 2018) provide alternatives to the traditional ‘point and click’ style of data analysis interactions which may provide benefits to non-traditional data users. One concern our participants had was that that mobile displays were not big enough for data analysis. Alternative interaction techniques could also decrease the amount of data needs to be shown directly to the user or optimising screen real estate to focus on things that really do require visualisation.

### **8.2.3 Accuracy and Trust of Sensing Technologies**

Across our studies we presented themes related to trust and device accuracy. In our studies we found that some users tended to either believe presented data without questioning it, which is in line with previous work (Hollis et al., 2018; Muir, 1987). However, we also found that some users questioned data when it did not meet their expectations of how the data should behave. In our exploratory study we found that users did not appear to be concerned with the scientific basis of the technology. In the study described in Chapter 5, we found that in some instances users did want to be aware of

the research the produced data was based on. Despite the inconsistency between our initial exploratory study, our qualitative results also suggested that users determine if a device is trustworthy based on their *Own Research* and *Inherent Trust in Technology*. Therefore we still suggest that it is important to be open about the basis on which the data is formed to ensure that both those with and without an interest in the scientific basis for a technology have adequate access to the information to engender trust.

An interrelated issue to trust is device accuracy, participants suggested they would be more inclined to trust a device that provided more accurate data. However, prior research suggests that users tend not to be able to determine how accurate a device is and a user's ability to know when a sensing device is providing accurate information (Huppert et al., 2019). We suggest based on our studies that designers and developers of PI sensing device should provide users with means of assessing device accuracy. Participants in our study suggested that the ability to *Compare and Correlate Sensor Results* against something that they can quantify (e.g. counting steps matches measured steps) or the ability to calibrate values against another trusted sensor, as possible means of doing so. Further research is required to determine how best to calibrate accuracy for values that is not so easily quantifiable (e.g. mental effort) between a user expectation and a sensing technology's output.

Design considerations related to accuracy may result in increased user trust in sensing technologies, based on our qualitative results which suggested that users would *Trust Values That Are More Accurate* and *Trust Values That Meet Expectations*. However, we found in our final study that data visualisations meeting prior expectations, the precision of the presented data, real-time or post-hoc visualisation, and prior knowledge of the device do not play a significant role in users' trust in the data produced by sensing technology. Rather we suggest, quantitatively, that initial confidence in a technology may be a determining factor that influences whether a person will trust the technology. Although we did not specifically seek to determine whether trustworthiness or accuracy of a technology is a factor in deciding whether to use it, prior work suggests that trustworthiness is a factor in continued usage of a device (Rupp et al., 2016). These considerations could have implications on the 'preparation' and 'selecting' phases defined by I. Li et al. (2010) and Epstein et al. (2014), respectively in their models of PI. Factors such as convenience, price, brand reputation could also be relevant to different degrees; Epstein et al. suggests that "tool selection can depend on features, aesthetics, and convenience, and choices can be limited by the tracker's mobile platform or budget." (Epstein et al., 2014). Future research could seek to understand what factors are relevant when deciding what device to use to track, and whether trust and



accuracy is a consideration by users.

## **8.3 Limitations**

### **8.3.1 Sampling & Selection Bias**

In both of our in-the-wild studies we had small sample sizes of 10 participants and nine participants, for the first and second study respectively. The participants in our second in-the-wild study were self-selected and therefore may not represent the entire population of people living with ME/CFS, CF, or ICF. In particular the symptoms of fatigue and the variation between people can be quite personal. The study does not account for the fact that people living with daily debilitating fatigue may not make use of the technologies as easily, or may not have the energy to be able to analyse the data. In all three of the studies participants self-selected to participate in the study, thus they are perhaps likely to be slightly more biased toward those who are willing to use new technologies. Further work may attempt to increase the number of participants, to increase the likelihood of encountering people with a larger variation of technological acceptance.

The majority of participants who took part in our final lab-based trust study were university-educated individuals and therefore may have better understandings of how data is graphed compared to the general population, which may impact how the graphed visualisations were interpreted. Again, increasing the sample size to include more participants from the general population to be more representative may produce alternative results.

### **8.3.2 Novelty Effect**

Participants in some of our studies suggested that part of the reason for using the devices was that it was new or novel. Our results therefore do not account for longer-term tracking and how once the novelty effect of using the technologies wears off may change users' views and what additional challenges they may encounter. Based on participant's desire for tracking longer-term trends, mentioned in Chapter 5, longer studies may provide additional interesting insights related to long-term tracking and considerations for when the effect of device novelty has worn off, and would likely be of interest to those tracking for chronic conditions.

## 8.4 Conclusion

This thesis aimed to provide a better understanding of the challenges that may be faced by the increasing availability and use of sensing technologies for the purpose of self-understanding. In doing so we have provided a number of design considerations that may be beneficial in developing PI tools that enable better experiences for users tracking data. Additionally, we have provided insights into the factors related to accuracy and trust of sensing technologies when used within in the context of personal informatics. Further research is required to better understand how these suggestions transfer into long term usage of PI tools and with other groups of users. However, we hope the research presented in this thesis provides useful insights to those developing PI tools, considerations for what users feel would be beneficial to their self-tracking experiences and challenges that still require further research to be addressed.

# Bibliography

- Allen, J. (2007, 3). Photoplethysmography and its application in clinical physiological measurement. *Physiological Measurement*, 28(3), R1. Retrieved from <https://iopscience.iop.org/article/10.1088/0967-3334/28/3/R01> doi: 10.1088/0967-3334/28/3/R01
- Alvarado, O., Heuer, H., Vanden Abeele, V., Breiter, A., & Verbert, K. (2020). Middle-Aged Video Consumers' Beliefs about Algorithmic Recommendations on YouTube. *Proceedings of the ACM on Human-Computer Interaction*, 4(CSCW2). Retrieved from <https://doi.org/10.1145/3415192> doi: 10.1145/3415192
- Angelakis, E., Lubar, J. F., Stathopoulou, S., & Kounios, J. (2004). Peak alpha frequency: An electroencephalographic measure of cognitive preparedness. *Clinical Neurophysiology*, 115(4), 887–897. doi: 10.1016/j.clinph.2003.11.034
- Antifakos, S., Kern, N., Schiele, B., & Schwaninger, A. (2005). Towards improving trust in context-aware systems by displaying system confidence. In *Acm international conference proceeding series* (Vol. 111, pp. 9–14). New York, New York, USA: ACM Press. Retrieved from <http://portal.acm.org/citation.cfm?doid=1085777.1085780> doi: 10.1145/1085777.1085780
- Armstrong, R. A. (2014, 9). When to use the Bonferroni correction. *Ophthalmic and Physiological Optics*, 34(5), 502–508. Retrieved from <https://onlinelibrary.wiley.com/doi/full/10.1111/opo.12131><https://onlinelibrary.wiley.com/doi/abs/10.1111/opo.12131><https://onlinelibrary.wiley.com/doi/10.1111/opo.12131> doi: 10.1111/OPO.12131
- Aroganam, G., Manivannan, N., & Harrison, D. (2019, 4). *Review on Wearable Technology Sensors Used in Consumer Sport Applications* (Vol. 19) (No. 9). NLM (Medline). doi: 10.3390/s19091983
- Ashford, R. (2014). Responsive and emotive wearables: Devices, bodies, data and communication. *Proceedings - International Symposium on Wearable Com-*

- puters, *ISWC*, 99–104. Retrieved from <http://dl.acm.org/citation.cfm?doid=2641248.2642731> doi: 10.1145/2641248.2642731
- Asmar, R., Khabouth, J., Topouchian, J., El Feghali, R., & Mattar, J. (2010, 2). Validation of three automatic devices for self-measurement of blood pressure according to the International Protocol: The Omron M3 Intellisense (HEM-7051-E), the Omron M2 Compact (HEM 7102-E), and the Omron R3-I Plus (HEM 6022-E). *Blood Pressure Monitoring*, 15(1), 49–54. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/20032779/> doi: 10.1097/MBP.0b013e3283354b11
- Ayobi, A., Marshall, P., Cox, A. L., & Chen, Y. (2017). Quantifying the body and caring for the mind: Self-tracking in multiple sclerosis. In *Conference on human factors in computing systems - proceedings* (Vol. 2017-May, pp. 6889–6901). Retrieved from <https://uclic.ucl.ac.uk/content/6-news-events-seminars/20170502-ucl-will-have-a-great-presence-at-chi-2017/ayobi.paper.pdf> doi: 10.1145/3025453.3025869
- Bentley, F., Tollmar, K., Stephenson, P., Levy, L., Jones, B., Robertson, S., . . . Wilson, J. (2013). Health mashups: Presenting statistical patterns between wellbeing data and context in natural language to promote behavior change. *ACM Transactions on Computer-Human Interaction*, 20(5), 30. Retrieved from <https://dl.acm.org/citation.cfm?doid=2533682.2503823> doi: 10.1145/2503823
- Berger, M. W. (2019, 8). *A wearable new technology moves brain monitoring from the lab to the real world — Penn Today*. Retrieved October 23, 2019, from [https://penntoday.upenn.edu/news/innovative-technology-wearable-portable-EEG-moves-brain-monitoring-from-lab-to-real-world?utm\\_source=Primary&utm\\_campaign=4a1f712116-EMAIL\\_CAMPAIGN\\_2019\\_07\\_23\\_11\\_35\\_COPY\\_01&utm\\_medium=email&utm\\_term=0\\_3777f2ca8f-4a1f712116](https://penntoday.upenn.edu/news/innovative-technology-wearable-portable-EEG-moves-brain-monitoring-from-lab-to-real-world?utm_source=Primary&utm_campaign=4a1f712116-EMAIL_CAMPAIGN_2019_07_23_11_35_COPY_01&utm_medium=email&utm_term=0_3777f2ca8f-4a1f712116)
- Bernstein, M., Hong, L., & Kairam, S. (2010). A torrent of tweets: managing information overload in online social streams. *CHI 2010 Workshop on Microblogging: What and How Can We Learn From It?*, 1–4. Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.188.7754>
- Bhayee, S., Tomaszewski, P., Lee, D. H., Moffat, G., Pino, L., Moreno, S., & Farb, N. A. (2016). Attentional and affective consequences of technology supported mindfulness training: A randomised, active control, efficacy trial. *BMC Psychology*, 4(1), 60. doi: 10.1186/s40359-016-0168-6
- Bogue, R. (2014). *Towards the trillion sensors market* (Vol. 34) (No. 2). Emerald Group Publishing Limited. doi: 10.1108/SR-12-2013-755
- Bolkhovskiy, J. B., Scully, C. G., & Chon, K. H. (2012). Statistical analysis of heart rate and heart rate variability monitoring through the use of smart phone cameras.

- In *Proceedings of the annual international conference of the IEEE Engineering in Medicine and Biology Society, EMBS* (pp. 1610–1613). doi: 10.1109/EMBC.2012.6346253
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77–101. Retrieved from <http://eprints.uwe.ac.uk/11735> doi: 10.1191/1478088706qp063oa
- Brunet, J. F., Dagenais, D., Therrien, M., Gartenberg, D., & Forest, G. (2017, 8). Validation of sleep-2-Peak: A smartphone application that can detect fatigue-related changes in reaction times during sleep deprivation. *Behavior Research Methods*, 49(4), 1460–1469. Retrieved from <https://link.springer.com/article/10.3758/s13428-016-0802-5> doi: 10.3758/s13428-016-0802-5
- Burke, L. E., Wang, J., & Sevvick, M. A. (2011, 1). Self-Monitoring in Weight Loss: A Systematic Review of the Literature. *Journal of the American Dietetic Association*, 111(1), 92–102. doi: 10.1016/j.jada.2010.10.008
- Canzian, L., & Musolesi, M. (2015). Trajectories of depression: Unobtrusive monitoring of depressive states by means of smartphone mobility traces analysis. In *UbiComp 2015 - proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing* (pp. 1293–1304). Retrieved from <http://dx.doi.org/10.1145/2750858.2805845> doi: 10.1145/2750858.2805845
- Chee-Keong Alfred, L., & Chong Chia, W. (2015, 4). Analysis of Single-Electrode EEG Rhythms Using MATLAB to Elicit Correlation with Cognitive Stress. *International Journal of Computer Theory and Engineering*, 7(2), 149–155. doi: 10.7763/ijcte.2015.v7.947
- Chi, E. H. (2008). Sensors and ubiquitous computing technologies in sports. In *Computers in sport* (pp. 249–267). Retrieved from [www.witpress.com](http://www.witpress.com), doi: 10.2495/978-1-84564-064-4/09
- Choe, E. K., Consolvo, S., Watson, N. F., & Kientz, J. A. (2011). *Opportunities for computing technologies to support healthy sleep behaviors*. Retrieved from <http://mdlabs.se/sleepcycle>. doi: 10.1145/1978942.1979395
- Choe, E. K., Lee, B., Kay, M., Pratt, W., & Kientz, J. A. (2015). SleepTight: Low-burden, self-monitoring technology for capturing and reflecting on sleep behaviors. In *UbiComp 2015 - proceedings of the 2015 ACM international joint conference on pervasive and ubiquitous computing* (pp. 121–132). Retrieved from <http://dx.doi.org/10.1145/2750858.2804266> doi: 10.1145/2750858.2804266
- Choe, E. K., Lee, B., & Schraefel, M. C. (2015, 7). Characterizing Visualization Insights from Quantified Selfers’ Personal Data Presentations. *IEEE Computer Graphics and Applications*, 35(4), 28–37. Retrieved from <http://ieeexplore.ieee.org/>

document/7106391/ doi: 10.1109/MCG.2015.51

- Choe, E. K., Lee, N. B., Lee, B., Pratt, W., & Kientz, J. A. (2014). Understanding quantified-selves' practices in collecting and exploring personal data. *Conference on Human Factors in Computing Systems - Proceedings*, 1143–1152. Retrieved from <http://dl.acm.org/citation.cfm?doid=2556288.2557372> <http://dl.acm.org/citation.cfm?id=2557372> doi: 10.1145/2556288.2557372
- Chuah, S. H. W., Rauschnabel, P. A., Krey, N., Nguyen, B., Ramayah, T., & Lade, S. (2016, 12). Wearable technologies: The role of usefulness and visibility in smartwatch adoption. *Computers in Human Behavior*, 65, 276–284. doi: 10.1016/j.chb.2016.07.047
- Cioroai, E., Buhnova, B., Kuhn, T., & Schneider, D. (2020). Building Trust in the Untrustable. *Proceedings - 2020 ACM/IEEE 42nd International Conference on Software Engineering: Software Engineering in Society, ICSE-SEIS 2020*, 21–24. Retrieved from <https://ieeexplore.ieee.org/abstract/document/9276510>
- Clayton, E. W. (2015, 3). *Beyond myalgic encephalomyelitis/chronic fatigue syndrome: An IOM report on redefining an illness* (Vol. 313) (No. 11). American Medical Association. Retrieved from <http://www.iom.edu/mecfs>. doi: 10.1001/jama.2015.1346
- Coates McCall, I., Lau, C., Minielly, N., & Illes, J. (2019). Owning Ethical Innovation: Claims about Commercial Wearable Brain Technologies. *Neuron*, 102(4), 728–731. Retrieved from <https://doi.org/10.1016/j.neuron.2019.03.026> doi: 10.1016/j.neuron.2019.03.026
- Crowley, K., Sliney, A., Pitt, I., & Murphy, D. (2010, 7). Evaluating a brain-computer interface to categorise human emotional response. In *Proceedings - 10th IEEE international conference on advanced learning technologies, icalt 2010* (pp. 276–278). IEEE. Retrieved from <http://ieeexplore.ieee.org/document/5571311/> doi: 10.1109/ICALT.2010.81
- Davies, T., Jones, S. L., & Kelly, R. M. (2019, 5). Patient perspectives on self-management technologies for chronic fatigue syndrome. In *Conference on human factors in computing systems - proceedings*. Association for Computing Machinery. doi: 10.1145/3290605.3300452
- Depper, A., & Howe, P. D. (2017, 1). Are we fit yet? English adolescent girls' experiences of health and fitness apps. *Health Sociology Review*, 26(1), 98–112. doi: 10.1080/14461242.2016.1196599
- Didžiokaitė, G., Saukko, P., & Greiffenhagen, C. (2018, 4). The mundane experience of everyday calorie trackers: Beyond the metaphor of Quantified Self. *New Media*

- and Society*, 20(4), 1470–1487. Retrieved from <http://journals.sagepub.com/doi/10.1177/1461444817698478> doi: 10.1177/1461444817698478
- Ding, X., Wei, S., Gui, X., Gu, N., & Zhang, P. (2021). Data Engagement Reconsidered: A Study of Automatic Stress Tracking Technology in Use. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Retrieved from <https://doi.org/10.1145/3411764.3445763><http://arxiv.org/abs/2101.05450> doi: 10.1145/3411764.3445763
- Doherty, K., & Doherty, G. (2019). Engagement in HCI: Conception, theory and measurement. *ACM Computing Surveys*, 51(5). Retrieved from <https://doi.org/10.1145/3234149> doi: 10.1145/3234149
- Dzindolet, M. T., Peterson, S. A., Pomranky, R. A., Pierce, L. G., & Beck, H. P. (2003, 6). The role of trust in automation reliance. *International Journal of Human Computer Studies*, 58(6), 697–718. doi: 10.1016/S1071-5819(03)00038-7
- Eikey, E. V., & Reddy, M. C. (2017). "It's definitely been a journey": A qualitative study on how women with eating disorders use weight loss apps. *Conference on Human Factors in Computing Systems - Proceedings, 2017-May*, 642–654. Retrieved from <http://dl.acm.org/citation.cfm?doid=3025453.3025591> doi: 10.1145/3025453.3025591
- Emotiv. (n.d.). *MyEmotiv*. Retrieved May 25, 2020, from <https://www.emotiv.com/myemotiv/>
- EMOTIV. (2017). *EMOTIV EPOC+ 14 Channel Mobile EEG - Emotiv*. Retrieved October 2, 2017, from <http://emotiv.com/product/emotiv-epoc-14-channel-mobile-eeg/>
- Eppler, M. J., & Mengis, J. (2008, 8). The Concept of Information Overload - A Review of Literature from Organization Science, Accounting, Marketing, MIS, and Related Disciplines (2004). In *Kommunikationsmanagement im wandel* (pp. 271–305). Gabler. Retrieved from [https://link.springer.com/chapter/10.1007/978-3-8349-9772-2\\_15](https://link.springer.com/chapter/10.1007/978-3-8349-9772-2_15) doi: 10.1007/978-3-8349-9772-2{\\\_}15
- Epstein, D. A., Caldeira, C., Figueiredo, M. C., Lu, X., Silva, L. M., Williams, L., ... Chen, Y. (2020, 12). Mapping and Taking Stock of the Personal Informatics Literature. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 4(4), 1–38. Retrieved from <https://dl.acm.org/doi/10.1145/3432231> doi: 10.1145/3432231
- Epstein, D. A., Caraway, M., Johnston, C., Ping, A., Fogarty, J., & Munson, S. A. (2016). Beyond abandonment to next steps: Understanding and designing for life after personal informatics tool use. In *Conference on human factors in computing systems - proceedings* (pp. 1109–1113). Retrieved from <http://dx.doi.org/>

10.1145/2858036.2858045 doi: 10.1145/2858036.2858045

- Epstein, D. A., Cordeiro, F., Bales, E., Fogarty, J., & Munson, S. A. (2014). Taming data complexity in lifelogs: Exploring visual cuts of personal informatics data. In *Proceedings of the conference on designing interactive systems: Processes, practices, methods, and techniques, dis* (pp. 667–676). New York, New York, USA: ACM Press. Retrieved from <http://dl.acm.org/citation.cfm?doid=2598510.2598558> doi: 10.1145/2598510.2598558
- Epstein, D. A., Lee, N. B., Kang, J. H., Agapie, E., Schroeder, J., Pina, L. R., ... Munson, S. A. (2017). Examining menstrual tracking to inform the design of personal informatics tools. In *Conference on human factors in computing systems - proceedings* (Vol. 2017-May, pp. 6876–6888). New York, New York, USA: ACM Press. Retrieved from <http://dl.acm.org/citation.cfm?doid=3025453.3025635> doi: 10.1145/3025453.3025635
- Epstein, D. A., Ping, A., Fogarty, J., & Munson, S. A. (2015). A lived informatics model of personal informatics. In *UbiComp 2015 - proceedings of the 2015 acm international joint conference on pervasive and ubiquitous computing* (pp. 731–742). New York, New York, USA: ACM Press. Retrieved from <http://dl.acm.org/citation.cfm?doid=2750858.2804250> doi: 10.1145/2750858.2804250
- Etiwy, M., Akhrass, Z., Gillinov, L., Alashi, A., Wang, R., Blackburn, G., ... Desai, M. Y. (2019). Accuracy of wearable heart rate monitors in cardiac rehabilitation. *Cardiovascular Diagnosis and Therapy*, 9(3), 262–271. Retrieved from <https://pubmed.ncbi.nlm.nih.gov/31275816/> doi: 10.21037/cdt.2019.04.08
- Falck, T., Baldus, H., Espina, J., & Klabunde, K. (2007, 6). Plug ’n play simplicity for wireless medical body sensors. *Mobile Networks and Applications*, 12(2-3), 143–153. Retrieved from <https://link.springer.com/article/10.1007/s11036-007-0016-2> doi: 10.1007/s11036-007-0016-2
- Fan, C., Forlizzi, J., & Dey, A. (2012a). Considerations for technology that support physical activity by older adults. *ASSETS’12 - Proceedings of the 14th International ACM SIGACCESS Conference on Computers and Accessibility*, 33–40. Retrieved from <http://dl.acm.org/citation.cfm?doid=2384916.2384923> doi: 10.1145/2384916.2384923
- Fan, C., Forlizzi, J., & Dey, A. K. (2012b). A spark of activity: Exploring informative art as visualization for physical activity. In *UbiComp’12 - proceedings of the 2012 acm conference on ubiquitous computing* (pp. 81–84). New York, New York, USA: ACM Press. Retrieved from <http://dl.acm.org/citation.cfm?doid=2370216.2370229> doi: 10.1145/2370216.2370229
- Felipe, S., Singh, A., Bradley, C., Williams, A. C., & Bianchi-Berthouze, N. (2015).



- Roles for Personal Informatics in chronic pain. In *Proceedings of the 2015 9th international conference on pervasive computing technologies for healthcare, pervasivehealth 2015* (pp. 161–168). Retrieved from [www.scavomed.com](http://www.scavomed.com) doi: 10.4108/icst.pervasivehealth.2015.259501
- Feustel, C., Aggarwal, S., Lee, B., & Wilcox, L. (2018, 9). People Like Me: Designing for Reflection on Aggregate Cohort Data in Personal Informatics Systems. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, 2(3). Retrieved from <https://doi.org/10.1145/3264917> doi: 10.1145/3264917
- Gharagozlou, F., Saraji, G. N., Mazloumi, A., Nahvi, A., Nasrabadi, A. M., Foroushani, A. R., ... Samavati, M. (2015, 12). Detecting driver mental fatigue based on EEG alpha power changes during simulated driving. *Iranian Journal of Public Health*, 44(12), 1693–1700. Retrieved from <http://ijph.tums.ac.ir>
- Gillinov, S., Etiwy, M., Wang, R., Blackburn, G., Phelan, D., Gillinov, A. M., ... Desai, M. Y. (2017). Variable accuracy of wearable heart rate monitors during aerobic exercise. *Medicine and Science in Sports and Exercise*, 49(8), 1697–1703. Retrieved from <http://www.acsm-msse.org> doi: 10.1249/MSS.0000000000001284
- Gimpel, H., Nißen, M., & Görlitz, R. A. (2013, 12). Quantifying the quantified self: A study on the motivation of patients to track their own health. *International Conference on Information Systems (ICIS 2013): Reshaping Society Through Information Systems Design*, 4, 3286–3301. Retrieved from <https://aisel.aisnet.org/icis2013/proceedings/HealthcareIS/3>
- Gulotta, R., Forlizzi, J., Yang, R., & Newman, M. W. (2016, 6). Fostering engagement with Personal informatics systems. In *Dis 2016 - proceedings of the 2016 acm conference on designing interactive systems: Fuse* (pp. 286–299). New York, New York, USA: ACM Press. Retrieved from <http://dl.acm.org/citation.cfm?doid=2901790.2901803> doi: 10.1145/2901790.2901803
- Haas, L. F. (2003, 1). Hans Berger (1873-1941), Richard Caton (1842-1926), and electroencephalography. *Journal of neurology, neurosurgery, and psychiatry*, 74(1), 9. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/12486257>  
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC1738204> doi: 10.1136/jnnp.74.1.9
- Häkikilä, J., Colley, A., Inget, V., Alhonsuo, M., & Rantakari, J. (2015). Exploring digital service concepts for healthy lifestyles. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 9186, pp. 470–480). Springer Verlag. Retrieved from [www.appannie.com](http://www.appannie.com) doi: 10.1007/978-3-319-20886-2\_{\\_}44
- Hall, D. L., & Llinas, J. (1997). An introduction to multisensor data fusion. *Proceedings*

- of the *IEEE*, 85(1), 6–23. doi: 10.1109/5.554205
- Harrison McKnight, D., Choudhury, V., & Kacmar, C. (2002, 12). The impact of initial consumer trust on intentions to transact with a web site: A trust building model. *Journal of Strategic Information Systems*, 11(3-4), 297–323. doi: 10.1016/S0963-8687(02)00020-3
- Hart, S. G., & Staveland, L. E. (1988, 1). Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. *Advances in Psychology*, 52(C), 139–183. Retrieved from <https://www.sciencedirect.com/science/article/pii/S0166411508623869> doi: 10.1016/S0166-4115(08)62386-9
- Hassib, M., Khamis, M., Schneegass, S., Shirazi, A. S., & Alt, F. (2016). Investigating user needs for bio-sensing and affective wearables. In *Conference on human factors in computing systems - proceedings* (Vol. 07-12-May-, pp. 1407–1414). Santa Clara, California, USA: ACM. doi: 10.1145/2851581.2892480
- Healio Endocrinetoday. (2020). *Coming soon to your smartwatch: CGM for healthy adults*. Retrieved May 3, 2020, from <https://www.healio.com/endocrinology/diabetes/news/print/endocrine-today/%7B619d7915-eec6-4d89-8ccc-ca23b4fe6a7f%7D/coming-soon-to-your-smartwatch-cgm-for-healthy-adults>
- Heintz, M., Govaerts, S., Law, E. L. C., Holzer, A., & Gillet, D. (2014). Pdot: Participatory design online tool. In *Conference on human factors in computing systems - proceedings* (pp. 2581–2586). New York, New York, USA: ACM Press. Retrieved from <http://dl.acm.org/citation.cfm?doid=2559206.2581266> doi: 10.1145/2559206.2581266
- Herculano-Houzel, S. (2002, 4). Do you know your brain? A survey on public neuroscience literacy at the closing of the decade of the brain. *Neuroscientist*, 8(2), 98–110. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/11954564> doi: 10.1177/107385840200800206
- Hines, T. M. (1998, 8). Comprehensive review of biorhythm theory. *Psychological Reports*, 83(1), 19–64. Retrieved from <http://journals.sagepub.com/doi/10.2466/pr0.1998.83.1.19> doi: 10.2466/pr0.1998.83.1.19
- Hollis, V., Pekurovsky, A., Wu, E., & Whittaker, S. (2018, 9). On Being Told How We Feel. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 2(3), 1–31. doi: 10.1145/3264924
- Huang, Y., Xu, J., Yu, B., & Shull, P. B. (2016, 7). Validity of FitBit, Jawbone UP, Nike+ and other wearable devices for level and stair walking. *Gait and Posture*, 48, 36–41. doi: 10.1016/j.gaitpost.2016.04.025
- Huppert, F., Kranz, M., & Hoelzl, G. (2019). Would i lie to you - would you notice?

- In Y. Yin, Y. Li, H. Gao, & J. Zhang (Eds.), *Lecture notes of the institute for computer sciences, social-informatics and telecommunications engineering, lnicst* (Vol. 290, pp. 230–243). Springer, Cham. doi: 10.1007/978-3-030-28468-8{\\_}17
- Jaimes, L. G., Murray, T., & Raij, A. (2013). Increasing trust in personal informatics tools. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 8014 LNCS, pp. 520–529). doi: 10.1007/978-3-642-39238-2{\\_}57
- Jerez-Fernandez, A., Angulo, A. N., & Oppenheimer, D. M. (2014, 2). Show Me the Numbers: Precision as a Cue to Others' Confidence. *Psychological Science*, 25(2), 633–635. Retrieved from <http://journals.sagepub.com/doi/10.1177/0956797613504301> doi: 10.1177/0956797613504301
- Jian, J.-Y., Bisantz, A. M., & Drury, C. G. (2000, 3). Foundations for an Empirically Determined Scale of Trust in Automated Systems. *International Journal of Cognitive Ergonomics*, 4(1), 53–71. Retrieved from <https://apps.dtic.mil/dtic/tr/fulltext/u2/a395339.pdf> doi: 10.1207/s15327566ijce0401{\\_}04
- Jones, S., & Kelly, R. (2016). Sensemaking Challenges in Personal Informatics and Self-Monitoring Systems. *ACM SIGCHI Conference on Human Factors in Computing Systems (CHI 2016)*, 2016–5. Retrieved from <http://opus.bath.ac.uk/>
- Jones, S. L., & Kelly, R. (2018, 3). Dealing With Information Overload in Multifaceted Personal Informatics Systems. *Human-Computer Interaction*, 33(1), 1–48. Retrieved from <https://www.tandfonline.com/doi/full/10.1080/07370024.2017.1302334> doi: 10.1080/07370024.2017.1302334
- Katz, D. S., Price, B. A., Holland, S., & Dalton, N. S. (2018, 4). Data, data everywhere, and still too hard to link: Insights From user interactions with diabetes apps. In *Conference on human factors in computing systems - proceedings* (Vol. 2018-April, pp. 1–12). New York, New York, USA: Association for Computing Machinery. Retrieved from <http://dl.acm.org/citation.cfm?doid=3173574.3174077> doi: 10.1145/3173574.3174077
- Kelley, C., Lee, B., & Wilcox, L. (2017). Self-tracking for mental wellness: Understanding expert perspectives and student experiences. In *Conference on human factors in computing systems - proceedings* (Vol. 2017-May, pp. 629–641). New York, New York, USA: ACM Press. Retrieved from <http://dl.acm.org/citation.cfm?doid=3025453.3025750> doi: 10.1145/3025453.3025750
- Kersten-van Dijk, E. T., Westerink, J. H., Beute, F., & IJsselstein, W. A. (2017). Personal Informatics, Self-Insight, and Behavior Change: A Critical Review of Current Literature. *Human-Computer Interaction*, 32(5-6), 268–296. doi: 10.1080/07370024.2016.1276456

- Kido, T. (2012). Self-tracking mindfulness incorporating a personal genome. *AAAI Spring Symposium - Technical Report, SS-12-05*, 31–36.
- Kim, D. J., Lee, Y., Rho, S., & Lim, Y. K. (2016). Design opportunities in three stages of relationship development between users and self-tracking devices. In *Conference on human factors in computing systems - proceedings* (pp. 699–703). Retrieved from <http://dl.acm.org/citation.cfm?doid=2858036.2858148> doi: 10.1145/2858036.2858148
- Kim, H. C. (2015). Acceptability engineering: The study of user acceptance of innovative technologies. *Journal of Applied Research and Technology*, 13(2), 230–237. Retrieved from [www.jart.ccadet.unam.mxwww.sciencedirect.com](http://www.jart.ccadet.unam.mxwww.sciencedirect.com) doi: 10.1016/j.jart.2015.06.001
- Kim, Y. S., Reinecke, K., & Hullman, J. (2018, 1). Data Through Others' Eyes: The Impact of Visualizing Others' Expectations on Visualization Interpretation. *IEEE Transactions on Visualization and Computer Graphics*, 24(1), 760–769. doi: 10.1109/TVCG.2017.2745240
- Klimesch, W. (1999). EEG alpha and theta oscillations reflect cognitive and memory performance: A review and analysis. *Brain Research Reviews*, 29(2-3), 169–195. Retrieved from <http://www.sciencedirect.com/science/article/pii/S0165017398000563> doi: 10.1016/S0165-0173(98)00056-3
- Koroleva, K., Krasnova, H., & Günther, O. (2010). 'STOP SPAMMING ME!' - Exploring information overload on Facebook. In *16th americas conference on information systems 2010, amcis 2010* (Vol. 4, pp. 2609–2618).
- Korotitsch, W. J., & Nelson-Gray, R. O. (1999). *An overview of self-monitoring research in assessment and treatment* (Vol. 11) (No. 4). American Psychological Association Inc. Retrieved from [/record/1999-01917-002](http://record/1999-01917-002) doi: 10.1037/1040-3590.11.4.415
- Koufaris, M., & Hampton-Sosa, W. (2004, 1). The development of initial trust in an online company by new customers. *Information and Management*, 41(3), 377–397. doi: 10.1016/j.im.2003.08.004
- Kulesza, T. (2012). An explanation-centric approach for personalizing intelligent agents. In *International conference on intelligent user interfaces, proceedings iui* (pp. 375–378). New York, New York, USA: ACM Press. Retrieved from <http://dl.acm.org/citation.cfm?doid=2166966.2167052> doi: 10.1145/2166966.2167052
- Kunze, K., Iwamura, M., Kise, K., Uchida, S., & Omachi, S. (2013, 10). Activity recognition for the mind: Toward a cognitive "Quantified Self". *Computer*, 46(10), 105–108. Retrieved from <http://ieeexplore.ieee.org/document/6649935/>

doi: 10.1109/MC.2013.339

- Laukkanen, R. M. T., & Virtanen, P. K. (1998). Heart rate monitors: State of the art. *Journal of Sports Sciences*, 16(sup1), 3–7. Retrieved from <https://www.tandfonline.com/action/journalInformation?journalCode=rjsp20> doi: 10.1080/026404198366920
- Lawson, S., Jamison-Powell, S., Garbett, A., Linehan, C., Kucharczyk, E., Verbaan, S., ... Morgan, K. (2013). Validating a mobile phone application for the everyday, unobtrusive, objective measurement of sleep. In *Conference on human factors in computing systems - proceedings* (pp. 2497–2506). New York, New York, USA: ACM Press. Retrieved from <http://dl.acm.org/citation.cfm?doid=2470654.2481345> doi: 10.1145/2470654.2481345
- Lawson, S., Kirman, B., Linehan, C., Feltwell, T., & Hopkins, L. (2015). Problematising upstream technology through speculative design: The case of quantified cats and dogs. In *Conference on human factors in computing systems - proceedings* (Vol. 2015-April, pp. 2663–2672). New York, New York, USA: ACM Press. Retrieved from <http://dl.acm.org/citation.cfm?doid=2702123.2702260> doi: 10.1145/2702123.2702260
- Lee, V. R. (2013). The Quantified Self (QS) movement and some emerging opportunities for the educational technology field. *Educational Technology*, 53(6), 39–42. Retrieved from [http://works.bepress.com/cgi/viewcontent.cgi?article=1015&context=victor\\_lee](http://works.bepress.com/cgi/viewcontent.cgi?article=1015&context=victor_lee)
- Li, G., Huang, S., Xu, W., Jiao, W., Jiang, Y., Gao, Z., & Zhang, J. (2020, 5). The impact of mental fatigue on brain activity: A comparative study both in resting state and task state using EEG. *BMC Neuroscience*, 21(1), 20. Retrieved from <https://bmcneurosci.biomedcentral.com/articles/10.1186/s12868-020-00569-1> doi: 10.1186/s12868-020-00569-1
- Li, I., Dey, A., & Forlizzi, J. (2010). A stage-based model of personal informatics systems. In *Conference on human factors in computing systems - proceedings* (Vol. 1, pp. 557–566). Retrieved from <http://dl.acm.org/citation.cfm?id=1753326.1753409> doi: 10.1145/1753326.1753409
- Li, I., Dey, A. K., & Forlizzi, J. (2011). Understanding my data, myself: Supporting self-reflection with ubicomp technologies. *UbiComp'11 - Proceedings of the 2011 ACM Conference on Ubiquitous Computing*, 405–414. Retrieved from <http://dl.acm.org/citation.cfm?doid=2030112.2030166> doi: 10.1145/2030112.2030166
- Li, I., Dey, A. K., & Forlizzi, J. (2012). Using context to reveal

- factors that affect physical activity. *ACM Transactions on Computer-Human Interaction*, 19(1). Retrieved from [http://delivery.acm.org/10.1145/2150000/2147790/a7-li.pdf?ip=138.38.103.148&id=2147790&acc=ACTIVESERVICE&key=BF07A2EE685417C5.85B475708465C551.4D4702B0C3E38B35.4D4702B0C3E38B35&CFID=808364577&CFTOKEN=23340492&\\_\\_acm\\_\\_=1506082944.f8530f31d52f8a4e13](http://delivery.acm.org/10.1145/2150000/2147790/a7-li.pdf?ip=138.38.103.148&id=2147790&acc=ACTIVESERVICE&key=BF07A2EE685417C5.85B475708465C551.4D4702B0C3E38B35.4D4702B0C3E38B35&CFID=808364577&CFTOKEN=23340492&__acm__=1506082944.f8530f31d52f8a4e13) doi: 10.1145/2147783.2147790
- Li, Y., Bahn, S., Nam, C. S., & Lee, J. (2014, 2). Effects of Luminosity Contrast and Stimulus Duration on User Performance and Preference in a P300-Based Brain-Computer Interface. *International Journal of Human-Computer Interaction*, 30(2), 151–163. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/10447318.2013.839903> doi: 10.1080/10447318.2013.839903
- Liang, Z., Nagata, Y., Martell, M. A. C., & Nishimura, T. (2016, 11). Nurturing wearable and mHealth technologies for self-care: Mindset, tool set and skill set. In *2016 IEEE 18th International Conference on e-Health Networking, Applications and Services, HealthCom 2016*. Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/HealthCom.2016.7749432
- Lim, B. Y. (2010). Improving trust in context-aware applications with intelligibility. In *Ubicomp'10 - proceedings of the 2010 ACM conference on ubiquitous computing* (pp. 477–480). New York, New York, USA: ACM Press. Retrieved from <http://portal.acm.org/citation.cfm?doid=1864431.1864491> doi: 10.1145/1864431.1864491
- Lin, J. J., Mamykina, L., Lindtner, S., Delajoux, G., & Strub, H. B. (2006). Fish'n'Steps: Encouraging physical activity with an interactive computer game. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 4206 LNCS, pp. 261–278). Springer, Berlin, Heidelberg. doi: 10.1007/11853565{\\_}16
- Looney, D., Kidmose, P., & Mandic, D. P. (2014). Ear-EEG: User-centered and wearable BCI. In C. Guger, B. Allison, & E. C. Leuthardt (Eds.), *Biosystems and biorobotics* (Vol. 6, pp. 41–50). Berlin, Heidelberg: Springer Berlin Heidelberg. doi: 10.1007/978-3-642-54707-2{\\_}5
- Lupton, D. (2014, 12). Self-tracking cultures: Towards a sociology of personal informatics. In *Proceedings of the 26th Australian computer-human interaction conference, OzCHI 2014* (pp. 77–86). New York, New York, USA: Association for Computing Machinery, Inc. Retrieved from <http://dl.acm.org/citation.cfm?doid=2686612.2686623> doi: 10.1145/2686612.2686623
- Lupton, D. (2017, 1). *Self-tracking, health and medicine* (Vol. 26) (No. 1). Routledge.

- doi: 10.1080/14461242.2016.1228149
- Lupton, D. (2020, 1). Data mattering and self-tracking: what can personal data do? *Continuum*, 34(1), 1–13. doi: 10.1080/10304312.2019.1691149
- MacLeod, C. M. (1991). Half a century of reseach on the stroop effect: An integrative review. *Psychological Bulletin*, 109(2), 163–203. Retrieved from /record/1991-14380-001 doi: 10.1037/0033-2909.109.2.163
- Macleod, H., Tang, A., & Carpendale, S. (2013). Towards Personal Informatics Tools for Chronic Illness Management. In *Grand 2013: Rnote proceedings of the grand nce 2013 meeting* (Vol. 11, pp. 3–6). Retrieved from <http://hcitang.org/papers/2013-grand2013-towards-personal-informatics-tools.pdf>
- Maitland, J., Sherwood, S., Barkhuus, L., Anderson, I., Hall, M., Brown, B., ... Muller, H. (2006, 11). Increasing the awareness of daily activity levels with pervasive computing. In *2006 pervasive health conference and workshops, pervasivehealth* (pp. 1–9). IEEE. Retrieved from <http://ieeexplore.ieee.org/document/4205158/> doi: 10.1109/PCTHEALTH.2006.361667
- Malhi, G. S., Hamilton, A., Morris, G., Mannie, Z., Das, P., & Outhred, T. (2017, 11). *The promise of digital mood tracking technologies: Are we heading on the right track?* (Vol. 20) (No. 4). BMJ Publishing Group. doi: 10.1136/eb-2017-102757
- Mamykina, L., Smaldone, A. M., & Bakken, S. R. (2015, 8). Adopting the sense-making perspective for chronic disease self-management. *Journal of Biomedical Informatics*, 56, 406–417. doi: 10.1016/j.jbi.2015.06.006
- Mann, S. (1997, 2). Wearable computing: A first step toward personal imaging. *Computer*, 30(2), 25–32. doi: 10.1109/2.566147
- Mann, S. (2004a). Continuous lifelong capture of personal experience with eyeTap. In *Carpe'04 - proceedings of the first acm workshop on continuous archival and retrieval of personal experiences* (pp. 1–21). New York, New York, USA: ACM Press. Retrieved from <http://portal.acm.org/citation.cfm?doid=1026653.1026654> doi: 10.1145/1026653.1026654
- Mann, S. (2004b). "Sousveillance". In *Proceedings of the 12th annual acm international conference on multimedia - multimedia '04* (p. 620). New York, New York, USA: Association for Computing Machinery (ACM). Retrieved from <http://portal.acm.org/citation.cfm?doid=1027527.1027673> doi: 10.1145/1027527.1027673
- Mann, S. (2018). *The Encyclopedia of Human-Computer Interaction, 2nd Ed. 23. Wearable Computing*. Retrieved June 28, 2020, from <https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/wearable-computing>

- Mann, S., Fung, J., & Moncrieff, E. (1999, 10). EyeTap technology for wireless electronic news gathering. *ACM SIGMOBILE Mobile Computing and Communications Review*, 3(4), 19–26. Retrieved from <https://dl.acm.org/doi/10.1145/584039.584044> doi: 10.1145/584039.584044
- Maskeliunas, R., Damasevicius, R., Martisius, I., & Vasiljevas, M. (2016). Consumer-grade EEG devices: Are they usable for control tasks? *PeerJ*, 2016(3), e1746. Retrieved from <https://peerj.com/articles/1746> doi: 10.7717/peerj.1746
- Matthews, M., Doherty, G., Sharry, J., & Fitzpatrick, C. (2008, 5). Mobile phone mood charting for adolescents. *British Journal of Guidance and Counselling*, 36(2), 113–129. Retrieved from <http://www.tandfonline.com/doi/abs/10.1080/03069880801926400> doi: 10.1080/03069880801926400
- Matthews, M., Murnane, E., & Snyder, J. (2017, 11). Quantifying the Changeable Self: The Role of Self-Tracking in Coming to Terms With and Managing Bipolar Disorder. *Human-Computer Interaction*, 32(5-6), 413–446. Retrieved from <https://www.tandfonline.com/doi/full/10.1080/07370024.2017.1294983> doi: 10.1080/07370024.2017.1294983
- McCall, B. (2018, 10). *Innovation awards finalist takes on Alzheimer's with 'Fitbit for the brain'*. Retrieved from <https://www.irishtimes.com/business/innovation/innovation-awards-finalist-takes-on-alzheimer-s-with-fitbit-for-the-brain-1.3664322>
- McFall, L. (2019, 1). Personalizing solidarity? The role of self-tracking in health insurance pricing. *Economy and Society*, 48(1), 52–76. doi: 10.1080/03085147.2019.1570707
- Meyer, A. N., Murphy, G. C., Zimmermann, T., & Fritz, T. (2017, 12). Design recommendations for self-monitoring in the workplace: Studies in software development. *Proceedings of the ACM on Human-Computer Interaction*, 1(CSCW), 1–24. Retrieved from <http://dl.acm.org/citation.cfm?doid=3171581.3134714> doi: 10.1145/3134714
- Meyer, J., Fortmann, J., Wasmann, M., & Heuten, W. (2015). Making lifelogging usable: Design guidelines for activity trackers. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 8936, pp. 323–334). Springer Verlag. doi: 10.1007/978-3-319-14442-9\_{-}39
- Meyer, J., Kazakova, A., Büsing, M., & Boll, S. (2016). Visualization of complex health data on mobile devices. In *Mmhealth 2016 - proceedings of the 2016 acm workshop on multimedia for personal health and health care, co-located with acm multimedia 2016* (pp. 31–34). Retrieved from <http://dx.doi.org/10.1145/>



2985766.2985774 doi: 10.1145/2985766.2985774

- Michaelis, J. R., Rupp, M. A., Kozachuk, J., Ho, B., Zapata-Ocampo, D., McConnell, D. S., & Smither, J. A. (2016). Describing the user experience of wearable fitness technology through online product reviews. In *Proceedings of the human factors and ergonomics society* (pp. 1072–1076). doi: 10.1177/1541931213601248
- Micklewright, D., St Clair Gibson, A., Gladwell, V., & Al Salman, A. (2017). Development and Validity of the Rating-of-Fatigue Scale. *Sports Medicine*, 47(11), 2375–2393. doi: 10.1007/s40279-017-0711-5
- Miele, S., & Shockley, R. I. (2013). The real-world use of big data. *IBM Global Business Services Saïd Business School at the University of Oxford*, 12(12), 15. Retrieved from <http://public.dhe.ibm.com/common/ssi/ecm/en/gbe03550usen/GBE03550USEN.PDF>
- Millais, P., Jones, S. L., & Kelly, R. (2018). Exploring data in virtual reality: Comparisons with 2d data visualizations. In *Conference on human factors in computing systems - proceedings* (Vol. 2018-April). Retrieved from <https://doi.org/10.1145/3170427.3188537> doi: 10.1145/3170427.3188537
- Mills, R. (2011). Researching Social News-Is reddit.com a mouthpiece for the 'Hive Mind', or a Collective Intelligence approach to Information Overload? In *Ethicomp 2011 proceedings* (pp. 1–16). Retrieved from [www.yourfreedom.hmg.gov.uk](http://www.yourfreedom.hmg.gov.uk)
- Mladenov, M., & Mock, M. (2009). A step counter service for java-enabled devices using a built-in accelerometer. In *Acm international conference proceeding series* (Vol. 385, pp. 1–5). doi: 10.1145/1554233.1554235
- Mousa, M., Zhang, X., & Claudel, C. (2016, 10). Flash Flood Detection in Urban Cities Using Ultrasonic and Infrared Sensors. *IEEE Sensors Journal*, 16(19), 7204–7216. doi: 10.1109/JSEN.2016.2592359
- Muir, B. M. (1987). Trust between humans and machines, and the design of decision aids. *International Journal of Man-Machine Studies*, 27(5-6), 527–539. doi: 10.1016/S0020-7373(87)80013-5
- Mukhopadhyay, S. C. (2015). *Wearable sensors for human activity monitoring: A review* (Vol. 15) (No. 3). Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/JSEN.2014.2370945
- Muller, M. J. (2002). Participatory Design: The Third Space in HCI. In *The human-computer interaction handbook: Fundamentals, evolving technologies and emerging applications* (p. 1051–1068). USA: L. Erlbaum Associates Inc.
- Müller-Putz, G. R., & Pfurtscheller, G. (2008, 1). Control of an electrical prosthesis with an SSVEP-based BCI. *IEEE Transactions on Biomedical Engineer-*

- ing, 55(1), 361–364. Retrieved from <http://ieeexplore.ieee.org/document/4360033/> doi: 10.1109/TBME.2007.897815
- Ndaro, N. Z., & Wang, S. Y. (2018). Effects of Fatigue Based on Electroencephalography Signal during Laparoscopic Surgical Simulation. *Minimally Invasive Surgery, 2018*. doi: 10.1155/2018/2389158
- NeuroSky. (n.d.). *EEG Algorithms — NeuroSky*. Retrieved April 11, 2017, from <http://neurosky.com/biosensors/eeg-sensor/algorithms/>
- Neurosky. (2017). *EEG Headsets — NeuroSky Store*. Retrieved October 2, 2017, from <http://neurosky.com/biosensors/eeg-sensor/biosensors/>
- NHS Choices. (n.d.). *Electroencephalogram (EEG) - NHS Choices*. Department of Health. Retrieved October 2, 2017, from <http://www.nhs.uk/conditions/EEG/Pages/Introduction.aspx>
- Nickerson, R. S. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175–220. Retrieved from [/record/2018-70006-003](#) doi: 10.1037/1089-2680.2.2.175
- Niknejad, N., Ismail, W. B., Mardani, A., Liao, H., & Ghani, I. (2020, 4). A comprehensive overview of smart wearables: The state of the art literature, recent advances, and future challenges. *Engineering Applications of Artificial Intelligence, 90*, 103529. doi: 10.1016/j.engappai.2020.103529
- Nikolic-Popovic, J., & Goubran, R. (2011). Measuring heart rate, breathing rate and skin conductance during exercise. In *Memea 2011 - 2011 ieee international symposium on medical measurements and applications, proceedings* (pp. 507–511). IEEE Computer Society. doi: 10.1109/MeMeA.2011.5966751
- Ocak, H. (2009). Automatic detection of epileptic seizures in EEG using discrete wavelet transform and approximate entropy. *Expert Systems with Applications, 36*(2 PART 1), 2027–2036. doi: 10.1016/j.eswa.2007.12.065
- O’Hara, K. (2012). Trust in social machines: The challenges. In *Aisb/iacap world congress 2012: Social computing, social cognition, social networks and multiagent systems, social turn - snamas 2012, part of alan turing year 2012* (pp. 54–59).
- Olsen, A. L. (2018, 2). Precise performance: Do citizens rely on numerical precision as a cue of confidence? *Journal of Behavioral Public Administration, 1*(1). doi: 10.30636/jbpa.11.19
- Olsson, T. (2014). Layers of user expectations of future technologies: An early framework. *Conference on Human Factors in Computing Systems - Proceedings, 1957–1962*. Retrieved from <http://delivery.acm.org/10.1145/2590000/2581225/p1957-olsson.pdf?ip=138.38.103.128&id=2581225&acc=ACTIVESERVICE&key=BF07A2EE685417C5>

- .85B475708465C551.4D4702B0C3E38B35.4D4702B0C3E38B35&\_\_acm\_\_=  
1528411189\_2d94a0dd92a799d617a2cc72d46f7ecc doi: 10.1145/  
2559206.2581225
- Oxford English Dictionary. (n.d.-a). *Informatics*. Retrieved September 13, 2017, from  
<http://www.oed.com/view/Entry/95567>
- Oxford English Dictionary. (n.d.-b). *Personal*. Retrieved September 13, 2017, from  
<http://www.oed.com/view/Entry/141482>
- O'Hara, K., Tuffield, M. M., & Shadbolt, N. (2008, 12). Lifelogging: Privacy and  
empowerment with memories for life. *Identity in the Information Society*, 1(1),  
155–172. doi: 10.1007/s12394-009-0008-4
- People.cn. (2019). *AI headbands tracking student attention levels suspended amidst  
online controversy - People's Daily Online*. Retrieved April 16, 2020, from  
<http://en.people.cn/n3/2019/1101/c90000-9628768.html>
- Pfeiffer, J., Von Entress-Fürsteneck, M., Urbach, N., & Buchwald, A. (2016, 6).  
Quantify-ME: Consumer acceptance of wearable self-tracking devices. In *24th  
european conference on information systems, ecis 2016*. Retrieved from [https://  
aisel.aisnet.org/ecis2016\\_rp/99](https://aisel.aisnet.org/ecis2016_rp/99)
- Picard, R. W., & Healey, J. (1997). Affective wearables. *Personal and Ubiquitous  
Computing*, 1(4), 231–240. doi: 10.1007/BF01682026
- Pink, S., Sumartojo, S., Lupton, D., & Heyes La Bond, C. (2017, 6). Mundane  
data: The routines, contingencies and accomplishments of digital living. *Big Data  
and Society*, 4(1), 205395171770092. Retrieved from [http://journals.sagepub  
.com/doi/10.1177/2053951717700924](http://journals.sagepub.com/doi/10.1177/2053951717700924) doi: 10.1177/2053951717700924
- Pomfret, M. P., & Medford, J. L. (2007). Affective domain: Journaling. In *From  
principles of learning to strategies for instruction-with workbook companion* (pp.  
339–357). Springer US. Retrieved from [https://link.springer.com/chapter/  
10.1007/978-0-387-71086-0\\_12](https://link.springer.com/chapter/10.1007/978-0-387-71086-0_12) doi: 10.1007/978-0-387-71086-0{\\\_}12
- Quantified Self. (2019). *What is Quantified Self? - Quantified Self*. Retrieved May 4,  
2020, from <https://qsinstitute.com/about/what-is-quantified-self/>
- Rahman, M., Karwowski, W., Fafrowicz, M., & Hancock, P. A. (2019, 6). *Neuroergo-  
nomics applications of electroencephalography in physical activities: A systematic  
review* (Vol. 13). Frontiers. Retrieved from [https://www.frontiersin.org/  
article/10.3389/fnhum.2019.00182/full](https://www.frontiersin.org/article/10.3389/fnhum.2019.00182/full) doi: 10.3389/fnhum.2019.00182
- Raj, S., Lee, J. M., Garrity, A., & Newman, M. W. (2019, 3). Clinical Data in Context.  
*Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Techno-  
logies*, 3(1), 1–20. Retrieved from <https://dl.acm.org/doi/10.1145/3314409>  
doi: 10.1145/3314409

- Rapp, A., & Cena, F. (2014). Self-monitoring and technology: Challenges and open issues in personal informatics. In *Lecture notes in computer science (including sub-series lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 8516 LNCS, pp. 613–622). Retrieved from [https://link.springer.com/content/pdf/10.1007%2F978-3-319-07509-9\\_58.pdf](https://link.springer.com/content/pdf/10.1007%2F978-3-319-07509-9_58.pdf) doi: 10.1007/978-3-319-07509-9{\\_}58
- Rapp, A., & Cena, F. (2016). Personal informatics for everyday life: How users without prior self-tracking experience engage with personal data. *International Journal of Human Computer Studies*, 94, 1–17. doi: 10.1016/j.ijhcs.2016.05.006
- Rapp, A., Marcengo, A., Buriano, L., Ruffo, G., Lai, M., & Cena, F. (2018, 4). Designing a personal informatics system for users without experience in self-tracking: a case study. *Behaviour and Information Technology*, 37(4), 335–366. Retrieved from <https://www.tandfonline.com/doi/full/10.1080/0144929X.2018.1436592> doi: 10.1080/0144929X.2018.1436592
- Raptis, D., Tselios, N., Kjeldskov, J., & Skov, M. B. (2013). Does size matter? Investigating the impact of mobile phone screen size on users’ perceived usability, effectiveness and efficiency. In *Mobilehci 2013 - proceedings of the 15th international conference on human-computer interaction with mobile devices and services* (pp. 127–136). New York, New York, USA: ACM Press. Retrieved from <http://dl.acm.org/citation.cfm?doid=2493190.2493204> doi: 10.1145/2493190.2493204
- Ravi, N., Dandekar, N., Mysore, P., & Littman, M. L. (2005). Activity recognition from accelerometer data. In *Proceedings of the national conference on artificial intelligence* (Vol. 3, pp. 1541–1546). Retrieved from [www.aaai.org](http://www.aaai.org)
- Ravichandran, R., Sien, S. W., Patel, S. N., Kientz, J. A., & Pina, L. R. (2017). Making sense of sleep sensors: How sleep sensing technologies support and undermine sleep health. *Conference on Human Factors in Computing Systems - Proceedings, 2017-May*, 6864–6875. doi: 10.1145/3025453.3025557
- Rebolledo-Mendez, G., Dunwell, I., Martínez-Mirón, E. A., Vargas-Cerdán, M. D., De Freitas, S., Liarokapis, F., & García-Gaona, A. R. (2009). Assessing neurosky’s usability to detect attention levels in an assessment exercise. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 5610 LNCS, pp. 149–158). doi: 10.1007/978-3-642-02574-7{\\_}17
- Renaud, P., & Blondin, J. P. (1997, 9). The stress of stroop performance: Physiological and emotional responses to color-word interference, task pacing, and pacing speed. *International Journal of Psychophysiology*, 27(2), 87–97. Re-

- trieved from <https://pubmed.ncbi.nlm.nih.gov/9342640/> doi: 10.1016/S0167-8760(97)00049-4
- Rivera-Pelayo, V., Fessler, A., Müller, L., & Pammer, V. (2017). Introducing mood self-tracking at work: Empirical insights from call centers. *ACM Transactions on Computer-Human Interaction*, 24(1). Retrieved from <http://dx.doi.org/10.1145/3014058> doi: 10.1145/3014058
- Rooksby, J., Rost, M., Morrison, A., & Chalmers, M. (2014). Personal tracking as lived informatics. *Conference on Human Factors in Computing Systems - Proceedings*, 1163–1172. Retrieved from <http://dl.acm.org/citation.cfm?doid=2556288.2557039> doi: 10.1145/2556288.2557039
- Ruiz-Garcia, L., Lunadei, L., Barreiro, P., & Robla, J. I. (2009, 6). A review of wireless sensor technologies and applications in agriculture and food industry: State of the art and current trends. *Sensors (Switzerland)*, 9(6), 4728–4750. Retrieved from <http://www.mdpi.com/1424-8220/9/6/4728> doi: 10.3390/s90604728
- Rupp, M. A., Michaelis, J. R., McConnell, D. S., & Smither, J. A. (2016). The impact of technological trust and self-determined motivation on intentions to use wearable fitness technology. In *Proceedings of the human factors and ergonomics society* (pp. 1433–1437). Human Factors and Ergonomics Society Inc. doi: 10.1177/1541931213601329
- Russell, J. A., Weiss, A., & Mendelsohn, G. A. (1989). Affect Grid: A Single-Item Scale of Pleasure and Arousal. *Journal of Personality and Social Psychology*, 57(3), 493–502. Retrieved from <http://psycnet.apa.org/fulltext/1990-00158-001.pdf> doi: 10.1037/0022-3514.57.3.493
- Sawyer, T. L. (2019, 10). *Fitbit: Was It Ready for the Marketplace?: Editor: Thomas H. Sawyer* (Vol. 90) (No. 8). Routledge. doi: 10.1080/07303084.2019.1649563
- Seneviratne, S., Hu, Y., Nguyen, T., Lan, G., Khalifa, S., Thilakarathna, K., ... Seneviratne, A. (2017, 10). *A Survey of Wearable Devices and Challenges* (Vol. 19) (No. 4). Institute of Electrical and Electronics Engineers Inc. doi: 10.1109/COMST.2017.2731979
- Shemyakina, N. V., Nagornova, Z., & Ponomarev, V. a. (2010). Independent component analysis for creative versus non-creative task performance. *Advances in Bio-medical Research*, 222–228. Retrieved from <http://www.wseas.us/e-library/conferences/2010/Cambridge/MABIPH/MABIPH-24.pdf>
- Shin, G., Feng, Y., Jarrahi, M. H., & Gafinowitz, N. (2019, 4). Beyond novelty effect: a mixed-methods exploration into the motivation for long-term activity tracker use. *JAMIA Open*, 2(1), 62–72. Retrieved from <https://academic.oup.com/jamiaopen/article/2/1/62/5230941> doi: 10.1093/jamiaopen/ooy048

- Shin, G., Jarrahi, M. H., Fei, Y., Karami, A., Gafnowitz, N., Byun, A., & Lu, X. (2019, 5). *Wearable activity trackers, accuracy, adoption, acceptance and health impact: A systematic literature review* (Vol. 93). Academic Press Inc. doi: 10.1016/j.jbi.2019.103153
- Söllner, M., Hoffmann, A., Hoffmann, H., Wacker, A., & Leimeister, J. M. (2014). Understanding the formation of trust in IT artifacts. In *Socio-technical design of ubiquitous computing systems* (Vol. 2, pp. 39–58). doi: 10.1007/978-3-319-05044-7\_{-}{-}3
- Son, C.-G. (2019, 6). Differential diagnosis between “chronic fatigue” and “chronic fatigue syndrome”. *Integrative Medicine Research*, 8(2), 89–91. Retrieved from <https://linkinghub.elsevier.com/retrieve/pii/S221342201930071X> doi: 10.1016/j.imr.2019.04.005
- Spencer, E. A., Brassey, J., & Mahtani, K. (2017). Recall bias. In *Catalogue of bias 2017*. Retrieved from <https://catalogofbias.org/biases/recall-bias/>
- Springer, A., & Whittaker, S. (2019). Progressive disclosure empirically motivated approaches to designing effective transparency. In *International conference on intelligent user interfaces, proceedings iui* (Vol. Part F1476, pp. 107–120). New York, NY, USA: Association for Computing Machinery. Retrieved from <https://doi.org/10.1145/3301275.3302322> doi: 10.1145/3301275.3302322
- Suresh, K. (2011, 1). *An overview of randomization techniques: An unbiased assessment of outcome in clinical research* (Vol. 4) (No. 1). Wolters Kluwer – Medknow Publications. Retrieved from [https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3136079/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3136079/?report=abstract) doi: 10.4103/0974-1208.82352
- Swan, M. (2012, 11). Sensor mania! the internet of things, wearable computing, objective metrics, and the quantified self 2.0. *Journal of Sensor and Actuator Networks*, 1(3), 217–253. Retrieved from <http://www.mdpi.com/2224-2708/1/3/217> doi: 10.3390/jsan1030217
- Thorup, C. B., Andreassen, J. J., Sørensen, E. E., Grønkjær, M., Dinesen, B. I., & Hansen, J. (2017, 3). Accuracy of a step counter during treadmill and daily life walking by healthy adults and patients with cardiac disease. *BMJ Open*, 7(3), e011742. Retrieved from <http://bmjopen.bmj.com/> doi: 10.1136/bmjopen-2016-011742
- Tollmar, K., Bentley, F., & Viedma, C. (2012). Mobile health mashups: Making sense of multiple streams of wellbeing and contextual data for presentation on a mobile device. In *2012 6th international conference on pervasive computing technologies for healthcare and workshops, pervasivehealth 2012* (pp. 65–72). doi:

- 10.4108/icst.pervasivehealth.2012.248698
- Tsai, C. C., Lee, G., Raab, F., Norman, G. J., Sohn, T., Griswold, W. G., & Patrick, K. (2007, 6). Usability and feasibility of PmEB: A mobile phone application for monitoring real time caloric balance. *Mobile Networks and Applications*, 12(2-3), 173–184. Retrieved from <http://link.springer.com/10.1007/s11036-007-0014-4> doi: 10.1007/s11036-007-0014-4
- Vaizman, Y., Ellis, K., Lanckriet, G., & Weibel, N. (2018, 4). Extrasensory app: Data collection in-the-wild with rich user interface to self-report behavior. In *Conference on human factors in computing systems - proceedings* (Vol. 2018-April). Association for Computing Machinery. doi: 10.1145/3173574.3174128
- Van De Laar, B., Gürkök, H., Plass-Oude Bos, D., Poel, M., & Nijholt, A. (2013, 6). Experiencing BCI control in a popular computer game. *IEEE Transactions on Computational Intelligence and AI in Games*, 5(2), 176–184. doi: 10.1109/TCIAIG.2013.2253778
- Vidal, J. J. (1973, 6). Toward direct brain-computer communication. *Annual review of biophysics and bioengineering*, 2(1), 157–180. Retrieved from <http://www.annualreviews.org/doi/abs/10.1146/annurev.bb.02.060173.001105> doi: 10.1146/annurev.bb.02.060173.001105
- Videnova, L. (2016). *Understanding Personal Informatics Use in a Professional Football Club: A Rapid Ethnography* (Tech. Rep.). Retrieved from [https://uclic.ucl.ac.uk/content/2-study/4-current-taught-course/1-distinction-projects/2-16/videnova\\_lilyana.pdf](https://uclic.ucl.ac.uk/content/2-study/4-current-taught-course/1-distinction-projects/2-16/videnova_lilyana.pdf)
- Wall Street Journal. (2019). *In Chinese Classrooms, AI Monitors Students' Every Move*. Retrieved April 16, 2020, from <https://www.wsj.com/video/under-ais-watchful-eye-china-wants-to-raise-smarter-students/C4294BAB-A76B-4569-8D09-32E9F2B62D19.html>
- Walsh, G., & Golbeck, J. (2014). StepCity: A preliminary investigation of a personal informatics-based social game on behavior change. In *Conference on human factors in computing systems - proceedings* (pp. 2371–2376). New York, New York, USA: ACM Press. Retrieved from <http://dl.acm.org/citation.cfm?doid=2559206.2581326> doi: 10.1145/2559206.2581326
- Wearable Technologies. (2019). *Apple Patent Hints at Glucose Monitoring Sensor*. Retrieved May 3, 2020, from <https://www.wearable-technologies.com/2019/04/apples-new-patents-hint-at-watch-that-detects-smell-to-monitor-glucose/>
- Wilson, M. L., Sharon, N., Maior, H. A., Midha, S., Craven, M. P., & Sharples, S. (2018). Mental workload as personal data: designing a cognitive activity tracker.

- 3rd Symposium on Computing and Mental Health*, 1–5. Retrieved from <http://eprints.nottingham.ac.uk/50630/>
- Windthorst, P., Mazurak, N., Kuske, M., Hipp, A., Giel, K. E., Enck, P., ... Teufel, M. (2017, 2). Heart rate variability biofeedback therapy and graded exercise training in management of chronic fatigue syndrome: An exploratory pilot study. *Journal of Psychosomatic Research*, 93, 6–13. doi: 10.1016/j.jpsychores.2016.11.014
- Wolpaw, J. R., & Wolpaw, E. W. (2012). *Brain-computer interfaces : principles and practice*. Oxford University Press.
- Yang, R., Newman, M. W., & Forlizzi, J. (2014). Making sustainability sustainable: Challenges in the design of eco-interaction technologies. In *Conference on human factors in computing systems - proceedings* (pp. 823–832). New York, New York, USA: Association for Computing Machinery. Retrieved from <http://dl.acm.org/citation.cfm?doid=2556288.2557380> doi: 10.1145/2556288.2557380
- Yang, R., Shin, E., Newman, M. W., & Ackerman, M. S. (2015, 9). When fitness trackers don't 'fit': End-user difficulties in the assessment of personal tracking device accuracy. In *Ubicomp 2015 - proceedings of the 2015 acm international joint conference on pervasive and ubiquitous computing* (pp. 623–634). Association for Computing Machinery, Inc. doi: 10.1145/2750858.2804269
- Yin, E., Zhou, Z., Jiang, J., Chen, F., Liu, Y., & Hu, D. (2013, 4). A novel hybrid BCI speller based on the incorporation of SSVEP into the P300 paradigm. *Journal of Neural Engineering*, 10(2), 026012. Retrieved from <http://stacks.iop.org/1741-2552/10/i=2/a=026012> doi: 10.1088/1741-2560/10/2/026012
- Yoon, H., Doh, Y. Y., Yi, M. Y., & Woo, W. (2014). A conceptual framework for augmented smart coach based on quantified holistic self. In *Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)* (Vol. 8530 LNCS, pp. 498–508). Springer Verlag. doi: 10.1007/978-3-319-07788-8{\\_}46
- Young-Ho Kim, Bongshin Lee, Arjun Srinivasan, & Eun Kyoung Choe. (2021). Data@Hand: Fostering Visual Exploration of Personal Data on Smartphones Leveraging Speech and Touch Interaction. *Proceedings of the 2021 CHI Conference on Human Factors in Computing Systems (CHI '21)*. Retrieved from <https://doi.org/10.1145/3411764.3445421> doi: 10.1145/3411764.3445421
- Yu, W., & Ingalls, T. (2011). Trails - An interactive web history visualization and tagging tool. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 6770 LNCS(PART 2), 77–86. doi: 10.1007/978-3-642-21708-1{\\_}10
- Zheng, Y. L., Ding, X. R., Poon, C. C. Y., Lo, B. P. L., Zhang, H., Zhou, X. L.,



- ... Zhang, Y. T. (2014). Unobtrusive sensing and wearable devices for health informatics. *IEEE Transactions on Biomedical Engineering*, 61(5), 1538–1554. doi: 10.1109/TBME.2014.2309951
- Zinn, M. L., Zinn, M. A., & Jason, L. A. (2016). qEEG / LORETA in Assessment of Neurocognitive Impairment in a Patient with Chronic Fatigue Syndrome: A Case Report. *Clinical Research: Open Access ( ISSN 2469-6714 )*, 2(1). doi: 10.16966/2469-6714.110

## Appendix A

### Fitbit For the Mind?

---

# Fitbit for the Mind?: An Exploratory Study of ‘Cognitive Personal Informatics’

**Cillian Dudley**

University of Bath  
Bath, BA2 7AY, UK  
c.dudley@bath.ac.uk

**Simon L. Jones**

University of Bath  
Bath, BA2 7AY, UK  
s.l.jones@bath.ac.uk

Permission to make digital or hard copies of part or all of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for third-party components of this work must be honored. For all other uses, contact the Owner/Author.

*CHI'18 Extended Abstracts*, April 21–26, 2018, Montreal, QC, Canada

© 2018 Copyright is held by the owner/author(s).

ACM ISBN 978-1-4503-5621-3/18/04.

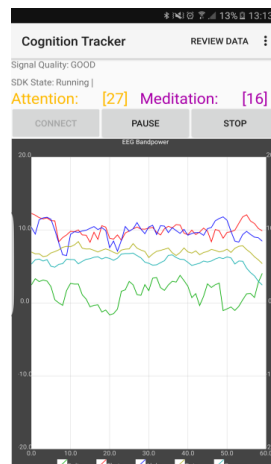
<https://doi.org/10.1145/3170427.3188530>

**Abstract**

Personal Informatics (PI) systems allow their users to collect data from a variety of sources for the purpose of extracting meaningful insights and making positive changes in their lives. Emerging consumer-grade Brain-Computer Interface (BCI)/EEG devices may provide an additional source of data for incorporating into PI systems. To explore users’ expectations for brain-related PI systems we provided participants with a consumer-grade BCI headset and prototype mobile application capable of visualizing and recording their brain waves. Participants were interviewed to assess expectations for this type of technology. Our work contributes an understanding of users’ various motivations for tracking brain activity data within a personal informatics system. We present our findings so far and discuss their implications for the design of a Cognitive Personal Informatics system, which we intend to deploy in a follow-up longitudinal field study.

**Author Keywords**

Personal Informatics; Brain Computer Interfaces; Quantified Self; EEG; Self-tracking; Health;



**Figure 1.** The Cognition Tracker Android application showing NeuroSky's eSense values (Attention 0-100, Meditation 0-100) and five EEG wave band powers (alpha, beta, gamma, delta, theta) in real-time

## ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous;

## Introduction

Personal Informatics (PI) and life-logging systems allow users to track data about their everyday activities and behaviors, and explore the collected data in order to uncover meaningful insights about themselves [3]. The variety of sources from which data can be collected is continuously expanding due to the emergence and availability of new wearable sensor technologies.

The recording and evaluation of EEG data is routinely used in clinical practice for detecting brain anomalies [6] and there is growing research into the use of EEG for controlling assistive technologies, e.g. prostheses [4]. However, the emergence of low-cost, consumer-grade EEG/BCI headsets from companies such as NeuroSky, Muse and Emotiv enables EEG recording devices to be obtained at reasonable cost for personal use. To date, very little attention has been paid to the potential role of EEG devices in the personal and lived informatics contexts described by Li et al. [3] and Rooksby et al. [7]. The willingness of users to capture physiological data about themselves has been shown by Hassib et al. [2].

Consumer-grade Brain Computer Interface (BCI) headsets, although currently in their infancy, may present an opportunity for the average consumer to track electroencephalogram (EEG) data, or 'brain data', offering users a figurative 'Fitbit for the mind'. This raises myriad questions about the use of EEG data in a personal informatics context. What value do users believe they can gain from recording EEG data? What

problems are people likely to experience when current consumer-grade BCI technologies are used for self-tracking? What HCI research challenges do we face in integrating BCI/EEG technologies with personal informatics systems?

In our ongoing work we seek to understand design opportunities, challenges, and technical, social and ethical implications for the "near-future" technology of 'Cognitive Personal Informatics' (CPI)—a class of tools that enables users to collect and analyze EEG data for the purpose of understanding and monitoring their brain activity..

## Exploratory Study Methodology

We have conducted an initial exploratory study to elicit people's initial reaction to an application capable of providing real-time feedback of EEG data as a tool for reflection rather than as an input/control device.

### *Cognition Tracker App*

The Cognition Tracker application (Figure 1) provided a simple line graph visualization of the five wave bands recorded by the headset; Alpha, Beta, Gamma, Delta, and Theta, plotted over a 60 second period and updated in real-time every second. In addition, two values representing meditation (mental calmness/relaxation) and attention (mental focus) derived from NeuroSky's proprietary algorithms [5], were displayed above the graph and also updated in real-time. The application served to give a practical demonstration of a BCI device acting as a real-time data tracker, rather than an input/control device.

Category for Use	No. of Participants
Improving self-understanding	8
Monitoring medical conditions	8
Optimizing behavior/performance	6
Hobbyist/technophile uses	5
Supplementing existing tracking technologies	4
Monitoring general health and wellbeing	3
Improving understanding of others	1

**Table 1.** Potential Categories of Use

#### Participants

16 participants (9 male, 7 female), aged 21-62 ( $M=30.08$ ,  $SD=10.69$ ), were recruited, via posting on the University of Bath's online noticeboard and by word of mouth. No specific requirements were needed for participation in the study. Participants had wide-ranging previous experiences with PI systems. Seven participants (P1, P3, P7, P11, P13, P14, P15) had previously used fitness trackers. One of these participants (P1) mentioned having used a range of tracking devices, including wearable fitness trackers and online services for 'life-logging' over a two-month period. When asked about technologies that they were already aware of for personal data tracking, participants provided examples for heart rate trackers, eye trackers, smart watches and wearable activity trackers, but none that focus on cognitive data.

#### Procedure

All participants were given a NeuroSky MindWave Mobile headset and a mobile device with the Cognition Tracker application (see Fig 1) installed to use for around 30 minutes. The participants were not given any specific instructions as to how they should use the system. Rather, participants were told they were free to use the application however they saw fit. Participants were given an initial introduction by the researcher, explaining how the application worked, what was presented on the display, and how to ensure the headset was transmitting correctly. Participants were then provided with help fitting the headset to ensure that it was positioned correctly, with a good quality signal connection, and that they knew how to begin recording data. Participants were then free to undertake any activity, e.g. going about their normal activities; working, reading, watching movies etc.,

whilst wearing the headset and having access to the Cognition Tracker application with the live data stream and historical data log. Participants took part in an interview shortly after using the headset and application. The interviewer asked participants about their initial experience and interactions with the system, and to discuss possible future uses and benefits of a system for recording EEG data. Participants were also asked if they had noticed anything interesting or intriguing in their data. Participants were prompted to identify any questions or hypotheses that they felt their EEG data might enable them to answer, and if there were other types of data they would consider combining with EEG data to learn more about themselves. Participants were asked if they had any concerns about recording their EEG data. All interview audio was transcribed and then inductively coded and thematically analyzed [1].

#### Interview Results and Discussion

##### *Why Use Cognitive Personal Informatics Systems?*

During the interviews participants were asked to consider the possible scenarios in which they felt EEG could be used. The primary purpose of this was to discover the meanings that people ascribe to the data and explore anticipated uses of the data. The categories of use suggested by participants can be seen in Table 1.

##### *What Insights Will Cognitive Personal Informatics Systems Provide?*

Table 1 summarizes a list of the metrics, cognitive processes and psychological states, which participants envisioned being able to monitor with the use of a personal informatics system. Some metrics show participants considering the devices as 'counters', i.e.

Metric / Process/ State	No. of Participants
Sleep rhythm	6
Focus attention	5
Stress / relaxation level	5
What my brain is doing / mental state	5
Concentration level	3
Disease progression	3
Productivity / efficiency	3
Cognitive load/ mental strain	2
Current mood (e.g. anger)	2
Biorhythm	1
Brain activation	1
Consciousness / fainting	1
Creativity	1
Depression	1
Liking / preference	1
Meditation state	1
Mind efficiency	1
Praying state	1
Procrastination	1

**Table 2.** Suggested metrics/states for tracking using CPI

solely producing quantitative data, in the same way that fitness devices are step and calorie counters. Cognitive tracking devices were viewed as quantitative ‘stress counters’, ‘cognitive load counters’, ‘brain activation counters’, and so on. The participant’s suggestions demonstrate their expectations that there is a broad range of meaningful, quantifiable values that can be obtained from a CPI system.

Other suggestions implied that participants believed there was rich, complex, qualitative data that could be captured. For example, suggesting that such a device could provide insights about “what my brain is doing” or “what I’m thinking” (P2), their “mental state” (P11), or “what’s actually going on when people are trying to be creative” (P9).

#### *Analyzing EEG Data in a CPI System*

Participants suggested a number of different types of analyses that they would like to perform (or for the system to perform automatically) on the basis of the data that the system was collecting. These analyses often implied that data would be: recorded frequently, in a wide variety of circumstances, over long periods of time, and fused with other forms of data to provide meaningful insights (see Table 2).

The most common type of analysis suggested by participants involved the comparison of brain wave data across different activities, to determine the effects of each activity on the user’s cognitive state. For example, P1 was interested in seeing if different activities lead to different patterns in their EEG: “Maybe doing sports, then reading a book then maybe have a call with a relative... I expect this will lead to different patterns in the EEG... It would be really interesting to see how

your brain behaves in certain situations.” Similarly, P7 was interested to see how his brain would respond to different activities: “it would be more out of curiosity just to see what happens to my brain when I do different things” and P14 wanted to find out about the effects of her environment on her mental state: “If I was to wear it for a longer period of time and maybe with like, in different environments, home environment, work environment, social, I could kinda see where I’m most comfortable maybe, most relaxed”.

Participants expressed interest in performing both inter-session and intra-session analyses. Inter-session analyses comprise comparing EEG data across distinct recording sessions, either for the same activity being performed at different times or in different settings, or comparing the data across recording sessions for different activities altogether. Intra-session analyses comprise a finer granularity of data being inspected in detail, for example drilling down into particular fluctuations in the EEG data within a single recording session and correlating them with particular external events to understand what effect they have on brain activity. P13 wanted to be able to switch between macro and micro level analyses, “zooming in” on interesting specific points within the data, e.g. a spike in attention values, and “zooming out” to see larger trends.

There was also interest from participants in being able to compare their own EEG data ‘to the norm’, e.g. “you could compare this EEG data with lots of other EEG datasets ... if my EEG data is comparable to the average healthy participant I would think OK, my brain, or the way I’m thinking, seems to be fine” (P4).

Data Sources
Blood Pressure
Heart Rate
Galvanic Skin Response
Browser History
Task/Activity Type
Physical Activity
Location/GPS
Diary
Gaze/Webcam/Video

**Table 3.** Participant suggested additional data types for combining with EEG



**Figure 2.** Cognition Tracker v2 Recording session

Participant 2 was keen to analyze data to uncover temporal patterns that reflected perceived variations in cognitive function, which they referred to as ‘biorhythms’; “[You could use this] to notice your biorhythm, to see when you’re most productive, to try to get the best out yourself, out of your mind...”

Four participants indicated that they would like to obtain summaries of their cognitive activity in the form of high-level information to accompany low-level EEG data, for example showing “summary statistics for each individual wave” (P13), daily values such as “maximum time spent concentrating and average amount of time spent concentrating” (P3), or statements such as “you have been very focused today” (P7), and “your EEG contains early warning signs that you might need to see a doctor” (P16)).

Analysis for the identification of triggers; external stimuli that activated a certain response in brain activity, was of importance to several participants, e.g. “I could potentially, by identifying that I find some things more stressful, find ways to try and limit that, to some extent” (P14).

Several participants reported experimenting with the Cognition Tracker tool during the study, deliberately altering their behavior and observing the output in order to try and understand how changes in their actions were manifested within the data. E.g. “I felt like I could separate certain waves by doing certain things... I wanted to just see how what I do has an effect on these values” (P13), “It felt like I could control the brain wave chart just by altering the way I was thinking, it seemed to correspond with something that was going on” (P16).

Often the identification of triggers implied the need for extremely rich contextual data collection, alongside EEG tracking. Participants wanted to be able to identify notable events within their EEG data and study the relationship with data that revealed contextual information about the event. For example, P4 suggested combining EEG with eye-tracking data: “...you could connect every visual impression with your biophysics and your brain activity, that would be interesting.”

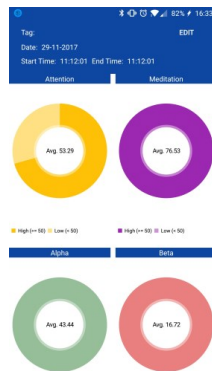
**Data Integration**  
Participants were asked what other data sources they would consider capturing alongside EEG data as well as the reason for doing so. They types of data suggested are shown in Table 3.

Data sources such as blood pressure, heart rate and galvanic skin response were suggested as being able to provide additional measures that might relate to cognitive activity (e.g. detecting stress). Whereas task, activity type, location/environment and diary records were suggested as a means of providing additional contextual data to support richer analysis of the EEG data (e.g. comparing emotional states between different locations or tasks).

**Future Work**  
Our ongoing work is expanding on this initial study by implementing Cognition Tracker v2 (Figures 2 – 4), a more feature-rich CPI tool that enables users to record and visualize their EEG data. While our work so far has elicited views and expectations about Cognitive Personal Informatics systems, based on a basic prototype intended to stimulate thinking about what the technology might be like, future work should aim to



**Figure 3.** Mean recorded eSense and waveband values for a single session.



**Figure 4.** Duration spent above/below a defined threshold value for each EEG measure.

capture participants' thoughts and experiences based on sustained use of a more fully-fledged CPI system. We plan to conduct a longitudinal study in which participants will be able to spend more time using the application to capture and review their data across different sessions, days, times, locations, and contexts.

Building on our findings from the exploratory study, the application will enable inter- and intra-session data analysis. Figure 3 shows the summary view of a single session and allows users to see their average state (i.e. meditation, attention) values, as well as normalized wave band data. Figure 4 shows information about the duration of the session spent in particular states (e.g. high or low attention states). Users will be able to compare these values between sessions by swiping through recording sessions.

Based on our participants' comments, Cognition Tracker v2 will allow users to tag their sessions with contextual information such as their activities at the time of the recording, how they were feeling, etc. to include in their reflection and analysis. The application will also allow for the integration of additional data sources suggested in Table 3, such as heart rate or GSR, which may enable more accurate measurements of cognitive activity than EEG alone.

We intend to evaluate users' experiences with the Cognition Tracker application in order to understand the role that CPI systems may play in improving self-understanding and mental wellbeing, and to inform the design of future PI systems that aim to integrate EEG and other emerging physiological sensor data.

## References

1. V Braun and V Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2: 77–101. <https://doi.org/10.1191/1478088706qp0630a>
2. Mariam Hassib, Mohamed Khamis, Stefan Schneegass, Ali Sahami Shirazi, and Florian Alt. 2016. Investigating User Needs for Bio-sensing and Affective Wearables. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems*, 1415–1422. <https://doi.org/10.1145/2851581.2892480>
3. Ian Li, Anind Dey, and Jodi Forlizzi. 2010. A stage-based model of personal informatics systems. In *Proceedings of the 28th international conference on Human factors in computing systems - CHI '10*, 557. <https://doi.org/10.1145/1753326.1753409>
4. Gernot R. Müller-Putz and Gert Pfurtscheller. 2008. Control of an electrical prosthesis with an SSVEP-based BCI. *IEEE Transactions on Biomedical Engineering* 55, 1: 361–364. <https://doi.org/10.1109/TBME.2007.897815>
5. Neurosky. EEG Algorithms. Retrieved April 11, 2017 from <http://neurosky.com/biosensors/eeg-sensor/algorithms>
6. Genaro Rebolledo-Mendez, Ian Dunwell, Erika A. Martínez-Mirín, María Dolores Vargas-Cerdán, Sara De Freitas, Fotis Liarokapis, and Alma R. García-Gaona. 2009. Assessing neurosky's usability to detect attention levels in an assessment exercise. In *Lecture Notes in Computer Science*, 149–158. [https://doi.org/10.1007/978-3-642-02574-7\\_17](https://doi.org/10.1007/978-3-642-02574-7_17)
7. John Rooksby, Mattias Rost, Alistair Morrison, and Matthew Chalmers Chalmers. 2014. Personal tracking as lived informatics. *Proceedings of the 32nd annual ACM conference on Human factors in computing systems - CHI '14*: 1163–1172. <https://doi.org/10.1145/2556288.2557039>



## Appendix B

### Study 1

## B.1 Participant Information and Consent Form

THE UNIVERSITY OF BATH, DEPARTMENT OF COMPUTER SCIENCE

## **PARTICIPANT INFORMATION AND CONSENT STATEMENT**

### *Cognition Tracker User Study I*

[ Contact Cillian Dudley [cd686@bath.ac.uk](mailto:cd686@bath.ac.uk) ]

#### *About this study*

During this study we will ask you to use and provide feedback on an Android application and its suitability for tracking cognitive data from a commercial brain-computer interface (BCI) headset.

BCI headsets have recently become commercially available, we are interested in understanding how they might help people to track data about themselves in a similar way to how people are using FitBit, and other wearables. We'd like to know if devices like the one you will be using can provide users with valuable insights into their daily activities and this study aims to explore whether BCI headsets could also be used in a similar manner.

This is an exploratory study and as such has no specific requirements other than the following:

We would like you to use the headset and application for around 30 minutes.

What you choose to do during this time is up to you but what you are doing should be noted down before you begin using the headset. Noting down your thoughts and ideas that come up throughout its use is encouraged.

The application provides users with real-time EEG feedback in the form of an Attention and Meditation value (from 0-100), and a graph of Alpha, Beta, Gamma, Delta and Theta band power which reflect your cognitive activity. For example, corresponding to state of awareness, emotional stress, daydreaming, focused concentration, relaxation, and creative inspiration.

**Should you have any questions or queries at any time please do not hesitate to contact the project team via Cillian Dudley, Department of Computer Science, University of Bath, Claverton Down, BA2 7AY. Email: [cd686@bath.ac.uk](mailto:cd686@bath.ac.uk)**

THE UNIVERSITY OF BATH, DEPARTMENT OF COMPUTER SCIENCE

**Confidentiality**

Any information that is obtained in connection with this study will be anonymised, and will be disclosed only with your permission, except as required by law. If you give us your permission by signing this document, we plan to discuss and publish the results to the wider research community. In any publication, information will be presented in such a way that you will not be able to be identified.

**Discomfort and risks**

This experiment is considered 'minimal risk'; the activities you will be asked to participate in are of no greater risk than those encountered in everyday life. The use of BCI headsets to facilitate recording of the brain's electrical activity should not cause any issues and as the headset is not certified for medical purposes the data collected should not create cause for concern. If you experience any discomfort at any point you are free to remove the headset and end the study.

**Your participation**

If you agree to participate by signing this form, your first involvement will be to use the BCI headset and application for 30 minutes at your own leisure, after which a followup interview will be conducted to gather your opinions and views of the application and headset, and also to gather feedback regarding your views on its potential uses. If you grant permission to do so, interviews will be recorded and later transcribed for analysis, however your participation does not require that the interview be recorded and an online survey can be supplied as an alternative to a face-to-face interview.

If you decide to participate, you are free to withdraw your consent and to discontinue your participation at any time without prejudice. This includes the right to have all experimental data concerning your participation destroyed. You will be given a copy of this form to keep.

**Should you have any questions or queries at any time please do not hesitate to contact the project team via Cillian Dudley, Department of Computer Science, University of Bath, Claverton Down, BA2 7AY. Email: [cd686@bath.ac.uk](mailto:cd686@bath.ac.uk)**

THE UNIVERSITY OF BATH, DEPARTMENT OF COMPUTER SCIENCE

**PARTICIPANT INFORMATION AND CONSENT STATEMENT (continued)**

*Cognition Tracker User Study I*

[ Contact Cillian Dudley [cd686@bath.ac.uk](mailto:cd686@bath.ac.uk) ]

**Participant No.:**

**Age:**

**Gender:**

**Occupation:**

**I agree to allow my anonymised data be used for this research?**

- ☐ Yes  
☐ No

**I agree to having a face-to-face interview, which will be recorded and transcribed?**

- ☐ Yes  
☐ No

**You are making a decision whether or not to participate in this study. Your signature indicates that having read the attached Participant Information and Consent Statement you have decided to take part in the study?**

.....  
Signature of Participant

.....  
Please PRINT name

.....  
Date

.....  
Signature of Researcher

.....  
Please PRINT name

.....  
Date

**Should you have any questions or queries at any time please do not hesitate to contact the project team via Cillian Dudley, Department of Computer Science, University of Bath, Claverton Down, BA2 7AY. Email: [cd686@bath.ac.uk](mailto:cd686@bath.ac.uk)**

THE UNIVERSITY OF BATH, DEPARTMENT OF COMPUTER SCIENCE

**WITHDRAW OF CONSENT BY PARTICIPANT**

***Cognition Tracker User Study 1***

I hereby wish to withdraw my consent for participation in the research study described above and understand that such withdrawal will not jeopardise any treatment, or my relationship with The University of Bath.

.....

Signature of Participant withdrawing

.....

Please PRINT name

.....

Date

**Should you have any questions or queries at any time please do not hesitate to contact the project team via Cillian Dudley, Department of Computer Science, University of Bath, Claverton Down, BA2 7AY. Email: [cd686@bath.ac.uk](mailto:cd686@bath.ac.uk)**

## B.2 Interview Schedule

### **Cognition Tracker User Study I - Questionnaire**

Any initial comments about the technology you've just used?

Did you encounter any issues using the application?

- What were they?

Did you encounter any issues using the hardware (Neurosky MindWave headset)?

- What were they?

Are you aware of, or have you had any experiences with, technologies that track data about you in ways that are similar to the application you have just used?

Have you used personal trackers/wearables before?

- Which ones?

What do you think the uses of a technology like this might be?

Who do you think would be interested in (or benefit from a technology like this)?

What are your thoughts about tracking EEG/brain/cognitive data?

What benefit do you think could be gained from recording EEG data on a regular basis?

Did anything about the recorded data stand out to (or intrigue) you?/Did you notice anything interesting when recording your EEG data?

What sort of questions do you think your EEG data could help you to answer?

Can you think of any other data you'd like to combine with EEG to learn more about yourself?

What task(s) were you doing during your recording session?

What other activities would you consider doing during future recording sessions?

What features about the application did you like in its current implementation?

Did you understand what the application was showing you?

What features would you suggest are added in future implementations?

If we were to give you a technology like this how do you think you would use it?



How long are you likely to spend recording your EEG data?

Where are you likely to be when willing to record your EEG data?

Do you have any concerns about recording EEG data?

## Appendix C

### Study 2

## C.1 Screening Questionnaire

THE UNIVERSITY OF BATH, DEPARTMENT OF COMPUTER SCIENCE

## ***Cognizance Tracker - Screening Questionnaire***

[ Contact Cillian Dudley [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk) ]

To ensure that we recruit and allocate participants suitably please provide the following details. A further sheet will follow with study details and consent should you be eligible. We may be unable to give details of ineligibility as this could affect the outcome of the study.

By filling in this form you agree that your anonymised details *may* be used in publication.

Only eligible participants will be included in the draw.

### **Personal Details**

Participant No.: .....

Name: .....

Age: .....

Gender: .....

Occupation: .....

Email Address: .....

Have you currently been diagnosed with a mental health or other medical condition?

☐ Yes ☐ No

### **Mobile Phone Details**

Make: .....

Model: .....

OS: ..... Version: .....

Reason for your interest in participating in this study:

.....  
.....  
.....

Should you have any questions or queries at any time please do not hesitate to contact the project team via Cillian Dudley, Department of Computer Science, University of Bath, Claverton Down, BA2 7AY. Email: [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk)

THE UNIVERSITY OF BATH, DEPARTMENT OF COMPUTER SCIENCE

**Prior Tracking Experience**

Prior tracking experience includes purposeful use of wearables or applications with the intention of understanding yourself better (e.g. on days where I do X, Y happens), changing behaviours (e.g. improving fitness by tracking steps), or goal achievement (e.g. calorie counting). Prior experience is not included just by passive use e.g. my phone, smartwatch (or other device) does this for me automatically but I don't interact with it.

**Have you previous tracked data about yourself for the purpose of self-understanding/self-improvement?**

☐ Yes ☐ No

**If yes, for how long have you tracked data?**

☐ < 1 month    ☐ < 3 months    ☐ < 6 months    ☐ < 12 months  
☐ 1-2 years    ☐ 2-5 years    ☐ 5+ years

**If yes, please list the devices or tools used:**

.....  
.....  
.....

**If yes, please describe the reasons for previous self-tracking:**

.....  
.....  
.....  
.....  
.....

**Should you have any questions or queries at any time please do not hesitate to contact the project team via Cillian Dudley, Department of Computer Science, University of Bath, Claverton Down, BA2 7AY. Email: [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk)**

**Mindfulness Experience**

**Do you have prior experience using mindfulness applications?**

For example, prior or current use mindfulness applications (e.g. Headspace, Calm, Insight Timer).

☐ Yes ☐ No

**If yes, for how long have you undertaken your mindfulness application use?**

☐ < 1 month   ☐ < 3 months   ☐ < 6 months   ☐ < 12 months  
☐ 1-2 years   ☐ 2-5 years   ☐ 5+ years

**Are you likely to undertake that practice during the duration of the study (likely dates July/August/September)?**

☐ Yes ☐ No

**Do you have prior experience using mindfulness practices (non-technological)?**

For example, prior/current mindfulness practices (e.g. Meditation)

☐ Yes ☐ No

**If yes, for how long have you undertaken your mindfulness practice?**

☐ < 1 month   ☐ < 3 months   ☐ < 6 months   ☐ < 12 months  
☐ 1-2 years   ☐ 2-5 years   ☐ 5+ years

**Are you likely to undertake that practice during the duration of the study (likely dates July/August/September)?**

☐ Yes ☐ No

**If yes, please list the devices or tools used:**

.....  
.....  
.....

**If yes, please describe your current/previous experience(s) with mindfulness:**

.....  
.....  
.....  
.....

**Should you have any questions or queries at any time please do not hesitate to contact the project team via Cillian Dudley, Department of Computer Science, University of Bath, Claverton Down, BA2 7AY. Email: [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk)**

## C.2 Participant Information and Consent Form

THE UNIVERSITY OF BATH, DEPARTMENT OF COMPUTER SCIENCE

## PARTICIPANT INFORMATION AND CONSENT STATEMENT

### *Cognition Tracker Study II*

[ Contact Cillian Dudley [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk) ]

#### About this study

During this study we will ask you to use and provide feedback on an Android application capable of visualising EEG, heart rate, breath rate, and mood data.

BCI headsets have recently become commercially available, we are interested in understanding how they might help people to track data about themselves in a similar way to how people are using FitBit, and other wearables. We'd like to know if devices like the one you will be using can provide users with valuable insights into their daily activities and this study aims to explore whether BCI headsets could also be used in a similar manner.

The application provides users with real-time EEG feedback in the form of Attention, Meditation, Appreciation, Alertness, Creativity, Cognitive Preparedness, Mental Effort, Familiarity, and eTensity with YinYang values. These values are recorded in real-time and are then provided as intra- and inter-session summaries. Descriptions for these values are provided within the applications 'Metric Info' section, accessible from the side menu.

The requested requirements of this study are:

We would like you to use the headset and application for around 30 minutes each day, as well as monitoring heart rate (via Fitbit) and breath rate (via Spire).

The EEG data can either be recorded on its own or with the addition of a mood log. Additional mood logs can be recorded without having to recorded a session with EEG data. At the end of recording EEG data you will be presented with a short workload assessment questionnaire which should be completed.

In addition to daily recording sessions using the application additional data will be recorded using Aware. This application will record the following information:

Application Use - What applications are being used.

Activity Recognition - What activities the application detects by your phone.

Location - Your location.

Communications - Logs non-identifiable communication events such as calls and messages, performed by or received. This does not record personal information, such as phone numbers or contact information

This data will not be published in an identifiable format, but rather used to discern trends in usage.

You will also be provided with a paper-based journal which can be used to record any additional insights, notes, or draw graphs that you find valuable.

To participate in the study is a requirement that currently do not suffer from, or be diagnosed with, any psychiatric condition.

**Should you have any questions or queries at any time please do not hesitate to contact the project team via Cillian Dudley, Department of Computer Science, University of Bath, Claverton Down, BA2 7AY. Email: [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk)**



**Confidentiality**

Any information that is obtained in connection with this study will be anonymised, and will be disclosed only with your permission, except as required by law. If you give us your permission by signing this document, we plan to discuss and publish the results to the wider research community. In any publication, information will be presented in such a way that you will not be able to be identified.

**Discomfort and risks**

This experiment is considered 'minimal risk'; the activities you will be asked to participate in are of no greater risk than those encountered in everyday life. The use of BCI headsets to facilitate recording of the brain's electrical activity should not cause any issues and as the headset is not certified for medical purposes the data collected should not create cause for concern. If you experience any discomfort at any point you are free to remove the headset.

**Your participation**

If you agree to participate by signing this form, your first involvement will be with the necessary hardware and software required to partake in the study. The researcher will explain all of the applications that are to be installed and how they and their associated software are used. It is requested that you use the Cognizance Tracker app and EEG headset at least once per day. This is necessary to be able to take part in the post-study interview.

After 3 weeks/21 days you will be asked to return to the hardware and undertake a one-to-one interview to gather to gather your opinions and views of using the Cognizance Tracker application. These interviews will be recorded and later transcribed for analysis. Assistance will be provided to uninstall any applications that you no longer wish to have on your phone.

If you decide to participate, you are free to withdraw your consent and to discontinue your participation at any time without prejudice. This includes the right to have all experimental data concerning your participation destroyed. You will be given a copy of this form to keep.

**Should you have any questions or queries at any time please do not hesitate to contact the project team via Cillian Dudley, Department of Computer Science, University of Bath, Claverton Down, BA2 7AY. Email: C.Dudley@bath.ac.uk**

THE UNIVERSITY OF BATH, DEPARTMENT OF COMPUTER SCIENCE

**PARTICIPANT INFORMATION AND CONSENT STATEMENT (continued)**

*Cognition Tracker User Study II*

[ Contact Cillian Dudley [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk) ]

**Participant No.:**

**Age:**

**Gender:**

**Occupation:**

**I agree to allow my anonymised data be used for research purposes?**

Yes ☐ No ☐

**I agree to having a face-to-face interview, which will be recorded and transcribed?**

Yes ☐ No ☐

**Have you currently been diagnosed with a mental health or other medical condition?**

Yes ☐ No ☐

**You are making a decision whether or not to participate in this study. Your signature indicates that having read the attached Participant Information and Consent Statement you have decided to take part in the study?**

.....

Signature of Participant

.....

Please PRINT name

.....

Date

.....

Signature of Researcher

.....

Please PRINT name

.....

Date

**Should you have any questions or queries at any time please do not hesitate to contact the project team via Cillian Dudley, Department of Computer Science, University of Bath, Claverton Down, BA2 7AY. Email: [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk)**

THE UNIVERSITY OF BATH, DEPARTMENT OF COMPUTER SCIENCE

**WITHDRAWAL OF CONSENT BY PARTICIPANT**

***Cognition Tracker User Study II***

I hereby wish to withdraw my consent for participation in the research study described above and understand that such withdrawal will not jeopardise any treatment, or my relationship with The University of Bath.

.....

Signature of Participant withdrawing

.....

Please PRINT name

.....

Date

**Should you have any questions or queries at any time please do not hesitate to contact the project team via Cillian Dudley, Department of Computer Science, University of Bath, Claverton Down, BA2 7AY. Email: [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk)**

### C.3 Interview Schedule

## Cognition Tracker User Study II - Questionnaire

### 1. INTRO

*In general...*

Do you have any initial comments about the technologies you've used over the past three weeks?

Did you encounter any issues using the application?

- What were they?

Did you encounter any issues using the hardware (Neurosky MindWave headset)?

- What were they?

### 2. PAST/PRESENT

*Thinking about your experience over the past three weeks with the device and headset...*

What are your thoughts about tracking EEG/brain/cognitive data?

What sort of questions do you think your EEG data could help you to answer?

- Has this changed since prior to taking part in this study? If so, how?

Do you feel you gained any benefit from recording your EEG data?

Did anything about the recorded data stand out to (or intrigue) you?/Did you notice anything interesting when recording your EEG data?

Did you discover any relationships between the data you were collecting? e.g. your concentration was affected by your mood.

What task(s) were you doing during your recording session?

What other activities would you consider doing during future recording sessions? / Were there any tasks you wanted to do but were unable to?

(List of tags, elaborate on why during that type of task, anything in particular looking to learn)

(And did/were you able to learn anything from the data about that?)

How did you decide on goal, metric, etc.

How did you use app to determine if you had reached X goal/value?

How many sessions did you log?

Why did you stop using the application device?

What would have made you continue using the device?

What features about the application did you like in its current implementation?

Did you understand what the application was showing you?

How accurate do you think the data was? Was it all equally accurate, were there bits that you trusted more than others?

- Do you think the data could still be useful given those inaccuracies?

Does the accuracy of the data change how you might use this technology?

Do you think the utility/usefulness would change over time with additional data collection?

How confident would you be making decisions based on the data you got?

Did you use all the features? Which most? Did you find the way the data was presented was useful?

Do you think the amount of information shown in the app was too much, just right, too little?

- What would you like to see less of? What you like it to focus on? - Why? - What would you like to see more of?

What were the most interesting, useful, important, surprising, obvious?

Did you trust the values that were being shown? Why (not)?

Are there any features or improvements that would have encouraged you to make more use of the application/device?

**Previous PI exp. only**

How did your usage of the Cognizance Tracker application compare to your previous experience with tracking applications e.g. [application used as per pre-screen]

**3. FUTURE**

*Thinking about the future...*

Can you think of any other data you'd like to combine with EEG to learn more about yourself?

What features would you suggest are added in future implementations

If we were to give you a technology like this how do you think you would use it?

How long are you likely to spend recording your EEG data?

Where are you likely to be when willing to record your EEG data?

What do you think the uses of a technology like this might be?

Who do you think would be interested in (or benefit from) a technology like this?

Do you have any concerns about recording EEG data?

Do you think there are any demerits from tracking EEG data?

[Privacy: What do you think an attacker could learn from your data?]

Do you think this type of emerging technology would have a positive or negative effect on your life?

#### **4. THINK ALOUD JOURNAL EXPLORATION**

## Appendix D

### Study 3



## D.1 Participant Information Sheet - Research Summary

## Research Summary



### **Understanding how people living with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome utilise data from emerging sensing technologies.**

The aim of this study is to determine how people with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (referred to for the purposes of this study as 'ME/CFS') might use new sensor technologies to gather data that enables them to reflect on their condition. The study aims to understand how people make use of new sensors which enable them to log data about themselves to promote helpful self-reflection.

To do this, participants with ME/CFS are being asked to use a smartphone with an application that integrates data from a Fitbit Versa and a NeuroSky MindWave headset over a three week period. The application will present this data in the form of graphs and charts of daily recording sessions as well as notifying users about data which they might find interesting based on insights derived from statistical analysis of their data. The application also provides the ability to further search the data using queries.

It is expected that this will take place between August and October 2019. All of the study activities are being undertaken as part of a PhD in Computer Science at the University of Bath. Participation in the study requires that you are diagnosed with ME or CFS, be over the age of 18, be willing to take part in all the three parts of the study (a pre-study briefing, three week period using the mobile app, and a post-study interview), and have sufficient internet access / mobile data to allow synchronisation of data.

At the end of the three week period using the devices you will be asked for your feedback by way of recorded interview. The interview questions will aim to determine if the application was suitable for your intended use. The post-study interview can be carried out in-person, or remotely (e.g. Skype, Messenger, telephone) as required by you.

If you choose to take part in the study, you will be provided with the required hardware to participate. A pre-study briefing will take place to help you get to grips with the hardware and to explain how the application is intended to be used. During the three week period you will be asked to make daily recording sessions and try to explore and make sense of the recorded data. After the three-week period of using the hardware and software, a post-study interview will be conducted to understand your experiences with the application and gather feedback about the way in which you've used the devices.

Participants will have the opportunity to try out an emerging technology and to take part in human-computer interaction research, so you may find it interesting to take part. The information that you and other participants provide in this project will help us to learn more about if and how people living with ME/CFS might be supported in future when using self-tracking or personal informatics tools.

One of the key aims of this project is to gather feedback from individuals with ME/CFS in such a way that does not negatively affect the condition, therefore we welcome conversation as to how to assist in minimising any symptomatic effect you might experience or be concerned might arise.

The devices used in this study have no known associated risks but there is a small risk that your interpretation of data presented to you may exacerbate your symptoms. Please always defer to your own knowledge and do not undertake non-normal daily activities. This is not a medical study, none of the devices used are medically certified, and the study does not aim to offer a treatment or cure for ME/CFS.

Participation in this research study is voluntary and you can withdraw at any stage of the project without being penalised or disadvantaged in any way. If you do decide to take part, you will be asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time and without giving a reason. Once any research papers, or the final thesis, are submitted for publication, you will no longer be able to withdraw the data.

Only the research team will have access to the information that you provide. All records will be treated as confidential. The written results may require use of direct quotations from some of the answers which you have given. Any quotations used will be anonymous and you may specify consent, or otherwise, for this on the Consent Form. Data produced as a result of the study will be stored in accordance with GDPR and the university's policies on data retention.

This project has been approved by the University of Bath, Research Ethics Approval Committee for Health (REACH) [reference: EP 18/19 028].

If you have questions you can contact the researchers, using the following details Cillian Dudley ([c.dudley@bath.ac.uk](mailto:c.dudley@bath.ac.uk)) or Dr Simon L. Jones ([s.l.jones@bath.ac.uk](mailto:s.l.jones@bath.ac.uk)) / +44 (0) 1225 385927).

If the researchers are unable to address your concerns or if you have a complaint you can contact the Chair of REACH, Professor James Betts, by email [j.betts@bath.ac.uk](mailto:j.betts@bath.ac.uk) or by telephone +44 (0) 1225 38 3448.

## D.2 Participant Information Sheet



## PARTICIPANT INFORMATION SHEET

Understanding how people living with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome utilise data from emerging sensing technologies.

**Researcher:** Cillian Dudley  
**Email:** [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk)

**Supervisor:** Dr Simon L. Jones  
**Email:** [S.L.Jones@bath.ac.uk](mailto:S.L.Jones@bath.ac.uk)  
**Telephone:** +44 (0) 1225 385927

This information sheet forms part of the process of informed consent. Before deciding to take part in the study it is important that you understand the purpose of the research and what your role will be. Please read the following information carefully. If required, you are permitted to discuss the briefing information with others. Please get in contact if you have any questions.

### 1. What is the purpose of the study?

The study aims to understand how people make use of new sensors which enable them to log data about themselves to promote helpful self-reflection. Tools and systems of this kind are known as “personal informatics” systems. An example of such a tool is Fitbit, which people use to track their steps and to determine if they have reached a specified target, or to encourage them to move more.

The aim of this study is to understand how people with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome (referred to for the purposes of this study as ‘ME/CFS’) might use new sensor technologies to gather data that enables them to reflect on their condition. To do this,

participants with ME/CFS are being asked to use a smartphone with an application that integrates data from a Fitbit Versa and a NeuroSky MindWave headset. These devices allow the collection of heart rate and brainwave data, respectively.



Fitbit Versa Watch  
(worn on wrist)



Neurosky MindWave  
(worn on head, as needed)

The brainwave data is converted into a variety of metrics or measures: Alertness, Appreciation, Attention, Cognitive Preparedness, Creativity, Familiarity, Meditation, Mental Effort, and Mood. These metrics are described as follows:

**Alertness:** Your level of alertness or vigilance. Higher values indicate a state of focus, lower values indicate a relaxing state of mind.

**Appreciation:** Your level of enjoyment or appreciation towards an external stimulus e.g. video, music, etc.

**Attention:** How focused you are at the moment, i.e. how much your attention is focused on a single thought or object.

**Cognitive Preparedness:** Your capacity for optimal cognitive performance on a relatively complex task i.e. your brain's capacity for higher level cognitive functions.

**Creativity:** Your underlying level of creative cognition. High values indicate stronger brainwave activity promoting innovation and creative

thinking.

**Familiarity:** Your learning progress while practicing a new skill. It gives an indication of the "learning curve" while acquiring a new skill.

**Meditation:** How calm or clear-minded you are at the moment.

**Mental Effort:** The mental workload experienced during a task. The harder your brain is working on a task, the higher the value.

**Emotion:** A combination of your emotional (pleasant, neutral or unpleasant) response and the intensity of that emotion toward an external stimulus.

The application will present this data in the form of graphs and charts of daily recording sessions as well as notifying users about data which they might find interesting based on insights derived from statistical analysis of their data. The application also provides the ability to further search the data using queries.

Your participation in the study will take three weeks, plus a pre- and post-study meeting. These meetings will last about 30 minutes and 1 hour respectively, but may vary to ensure ample time to gather your feedback and answer any questions you have.

It is expected that this will take place between August and October 2019. All of the study activities are being undertaken as part of a PhD in Computer Science at the University of Bath.

## **2. Why have I been selected to take part?**

To be eligible to take part in this study the participants must exhibit the following criteria:

✓ Be diagnosed with Myalgic Encephalomyelitis or Chronic Fatigue Syndrome



- ✓ Be over the age of 18
- ✓ Be willing to take part in the three parts of the study, consisting of:
  - A pre-study briefing (~30 minutes to 1 hour)
  - A three week period during August 2019 and October 2019, using the provided hardware and smartphone application.
  - A post-study interview of roughly 1 hour.
- ✓ Have internet access, or sufficient data on their mobile plan to allow synchronising data.

You have been selected as you fulfill these criteria and have shown an interest in participating in the study. We are expecting there to be 20 people participating in the research project.

### **3. Do I have to take part?**

Participation in this research study is voluntary. You can withdraw at any stage of the project without being penalised or disadvantaged in any way. One of the key aims of this project is to gather feedback from individuals with ME/CFS in such a way that does not negatively affect the condition, therefore we welcome conversation as to how to assist in minimising any symptomatic effect you might experience or be concerned might arise.

You can decide whether or not to take part. If you do decide to take part, you will be asked to sign a consent form. If you decide to take part, you are still free to withdraw at any time and without giving a reason. Once any research papers, or the final thesis, are submitted for publication, you will no longer be able to withdraw the data.

#### **4. What will happen if I take part?**

Over a three week period during August 2019 and October 2019 you will be asked to use a combination of devices (Fitbit and NeuroSky headset) and an Android mobile application. If you do not own an Android phone, one will be provided for the duration of the study to be used with the application.

You will be asked for your feedback at the end of the three week period by way of a recorded interview.

The interview questions will aim to determine if the application was suitable for your intended use, e.g. how well did it facilitate activity pacing?, as well as more general questions about how you used the application and hardware itself.

The post-study interview can be carried out in-person, or remotely (e.g. Skype, Messenger, telephone) as required by you.

The study will require no travel on your part and can be conducted in any place that you feel suitable. Any hardware will be delivered to you and collected.

#### **5. What do I have to do?**

If you choose to take part in the study, you will be provided hardware as described above. A pre-study briefing will take place to help you get to grips with the hardware and to explain how the application is intended to be used. You will be asked to install the following software on your phone: Cognizance Tracker, Fitbit, and AWARE (these applications are described in the following section) for the period of the study. Help will be provided to install and uninstall the required applications, once the study period is over. During the three-week period of using the hardware and software, you are free to use the devices as you wish but we request the following:

- Attempt to make a daily data recording session of about 30 minutes using the Cognizance Tracker application and NeuroSky MindWave headset.
- Make use of the Cognizance Tracker Insight Feed, and other application abilities to investigate the recorded data. Insight Feed items should be rated based on how interesting you find them, how much you trust what the item is telling you, how understandable the item is, and the utility or usefulness of the item.

After the three-week period of using the hardware and software, a post-study interview will be conducted to understand your experiences with the application and gather feedback about the way in which you've used the devices.

We welcome any discussion from you as to how the process could be improved or be made more achievable for you. You can withdraw from participation at any point.

#### **a. What do all these applications do?**

The following describes the four applications involved in this study and their purposes.

**Cognizance Tracker:** This is the main smartphone application used as part of the research. Cognizance Tracker takes data from all of the devices that you will be using. It analyses and presents the data to you to provide insights about your behaviours and activities.

The Cognizance Tracker application is also used for recording brainwave signals using the NeuroSky headset. This application works in combination with the NeuroSky headset as well as the 'companion' application for Fitbit.

**Fitbit:** Fitbit is the companion application for the Fitbit smartwatch. The Fitbit device's sensors collect a variety of data such as steps, heart rate and sleep. The heart rate data from Fitbit is integrated into the Cognizance tracker application.

<https://www.fitbit.com/>

Data recorded by the Fitbit application will automatically synchronize data over time into Cognizance Tracker for recorded entries. You are not required to actively engage with the Fitbit application unless you wish to do so, some interaction may be required to ensure that data is being recorded by them.

**AWARE** - AWARE is an app used by researchers to collect additional data that may be useful for their research. As part of this research the following data will be collected via the AWARE app:

- Application Use: What applications are being used on your phone and when.
- Activity Recognition: What activities the application detects based on your phone.
- Location: Your location as determined by your mobile phone's GPS or network service.
- Screen: The screen statuses, such as turning on and off, locked and unlocked.

It is hoped that this additional data may provide contextual information to Cognizance Tracker's use. You do not need to actively engage the AWARE app once installed.

#### **6. What are the possible advantages and benefits of taking part?**

Participants will have the opportunity to try out an emerging technology and to take part in human-computer interaction research, so you may find it interesting to take part. The information that you and other participants provide in this project will help us to learn more about if and how people living with ME/CFS might be supported in future when using self-tracking or personal informatics tools.

This is not a medical study, none of the devices used are medically certified, and the study does not aim to offer a treatment or cure for ME/CFS.

#### **7. What are the possible disadvantages and risks of taking part?**

None of the devices you will be asked to use have any known associated risks. However it may be that through the use of the devices, for example for pacing activities, that you interpret the data in such a way that suggests that you have more energy than you feel you do. You should not take any risks that you would not usually take in daily life and stick to your own feelings rather than doing what the device might suggest if you do not feel able. The main priority is your health, and should remain so during the study. You should not take any additional risks that you would consider being outside your normal day-to-day activities.

The questionnaire or actions required to complete the study tasks may be upsetting emotionally or physically draining to complete. The kind of questions asked will include questions about your symptoms, how you manage them, previous experiences self-tracking tools, and how you used the tools provided as part of this study. If at any stage you feel you've had enough of the questions or begin to experience adverse symptomatic effects on your ME/CFS condition, you are advised to stop. Doing the study remotely allows you to return to the task after a break,

complete it in smaller chunks, or ask others for help in completing it. As above, the main priority is your health and you can stop at any point.

The researcher is always available as a point of contact if you have any concerns.

**8. Who will have access to the information that I provide?**

Only the research team will have access to the information that you provide. All records will be treated as confidential.

**9. What will happen to the data when the research study stops?**

Any information that you provide will be handled in confidence and kept securely. No personal contact details will be presented in the final written dissertation, reports or publications. The research team will be the only people who will have access to the data during the study. Data will be analysed on computers which are password-protected, and written drafts of the research will also be stored securely in this way. The written results may require use of direct quotations from some of the answers which you have given. Any quotations used will be anonymous and you may specify consent, or otherwise, for this on the Consent Form.

**10. What should I do if I want to take part?**

If you wish to take part, please reply to this email or send a new email with the completed Informed Consent Form to the following address: [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk). An appropriate start date will be discussed and confirmed with you before commencing.

**11. What will happen to the results of the research study?**

All data collected during the project including personal, identifiable data will be treated as confidential and kept in a locked cabinet in a locked room or on a password protected file on the University of Bath's secure server (X drive). This storage of data will be done in accordance with GDPR. Recorded data will not be kept for any longer than 10 years.

Your name or other identifying information will not be disclosed in any presentation or publication of the research.

Once this project is completed, other researchers at the University of Bath may conduct related research projects which would benefit from the use of the data that you have provided. Further use of your data will only occur with your consent and the University of Bath's approval where data will continue to be stored in accordance with GDPR. So again, your name or other identifying information will not be disclosed in any presentation or publication of the research.

**12. What will happen if I do not want to carry on with the study?**

You are free to withdraw from the study without an explanation or consequence to yourself. If you wish to do so before completing all parts of the study, you can inform one of the researchers in person, email or telephone.

If, for any reason, you wish to withdraw your data please contact an identified researcher within two weeks of your participation. After this date it may not be possible to withdraw your data as some results may have been published. Your individual results however will not be identifiable in any way in any presentation or publication.

**13. Who has reviewed the study?**

This project has been approved by the University of Bath, Research Ethics Approval Committee for Health (REACH) [reference: EP 18/19 028].

**14. What if there is a problem?**

If you have a concern about any aspect of the project, you should ask to speak to the researchers (noted at the top of this document), who will do their best to answer any questions.

If they are unable to resolve your concern, or you wish to make a complaint regarding the project, please contact the Chair of the Research Ethics Approval Committee for Health.

Professor James Betts

**Email:** [j.betts@bath.ac.uk](mailto:j.betts@bath.ac.uk)

**Tel:** +44 (0) 1225 38 3448

Thank you very much for your time.



## D.3 Demographic Questionnaire



## Demographic Questionnaire

Understanding how people living with Myalgic Encephalomyelitis/Chronic Fatigue Syndrome utilise data from emerging sensing technologies.

You can fill in this form however you would like, using your computer, laptop or device. Alternatively, we will also accept scanned hand-written forms, or this questionnaire can be completed online at the following address: [http://bit.ly/CFS\\_ME\\_Questionnaire](http://bit.ly/CFS_ME_Questionnaire)

Please complete and return the form to [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk) with the consent form.

### 1. Personal Details

Name	
Age	
Gender	
Email Address	
Telephone Number	
Have you currently been diagnosed with Myalgic Encephalomyelitis and/or Chronic Fatigue Syndrome?	Yes / No (Please indicate).

## 2. Mobile Phone Details

Make	
Model	
Operating System	
Version (if known)	

## 3. Your reason for your interest in participating in this study

--

## 4. Prior Personal Informatics / Self-Tracking Experience

Prior tracking experience includes purposeful use of wearables or applications with the intention of understanding yourself better (e.g. on days where I do X, Y happens), changing behaviours (e.g. improving fitness by tracking steps), or goal achievement (e.g. calorie counting). Prior experience is not included just by passive use e.g. my phone, smartwatch (or other device) does this for me automatically but I don't interact with it.

Have you previous tracked data about yourself for the purpose of self-understanding/self-improvement?	Yes/No (Please indicate.)
---	------------------------------

<p>If yes, for how long have you tracked data?</p>	<div> <input type="checkbox"/> &lt; 1 month             <input type="checkbox"/> &lt; 3 months         </div> <div> <input type="checkbox"/> &lt; 6 months             <input type="checkbox"/> &lt; 12 months         </div> <div> <input type="checkbox"/> 1-2 years             <input type="checkbox"/> 2-5 years         </div> <div> <input type="checkbox"/> 5+ years         </div> <p>(Please indicate.)</p>
<p>If yes, please describe the reasons for previous self-tracking:</p>	
<div></div>	
<p>If yes, please list the devices or tools used:</p>	
<div></div>	

## D.4 Interview Schedule

## Interview Schedule

Understanding how people living with Chronic Fatigue Syndrome/Myalgic Encephalomyelitis utilise data from emerging sensing technologies.

### 1. INTRODUCTION

*In general...*

Do you have any initial comments about the technologies you've used over the past three weeks?

Did you encounter any issues using the Cognizance Tracker application?

- What were they?

Did you encounter any issues using the devices?

- What were they?

### 2. PAST/PRESENT

*Thinking about your experience over the past three weeks with the devices...*

What are your thoughts about tracking EEG/brain data?

What sort of questions do you think your EEG data could help you to answer?

- Has this changed since prior to taking part in this study? If so, how?

Do you feel you gained any benefit from recording your EEG data (in relation to ME/CFS)?

**Yes:** How does this compare to the benefit gained from recording your heart rate?

**No:** Do you think you gained any benefit from any of the data collected?

Did anything about the recorded data stand out to (or intrigue) you?

Did you discover any relationships between the data you were collecting?

**Yes:** How did you determine this relationship?

What task(s) were you doing during your recording sessions?

What other activities would you consider doing during future recording sessions?



Were there any tasks you wanted to do but were unable to?

x- How did you use app to determine if you had reached a goal/value?

Did you use the app to set targets or determine if you had reached certain values for any of the data?

Yes: How did you do that?

No: Do you think that is something you would be able to do with the app? Would it be beneficial? What would you need to be able to do it?

x - How many sessions did you log?

**If not used every day (< 21):**

Why did you stop using the application device?

What would have made you continue using the device?

**If not application related reasons**

Are there any features or improvements that would have encouraged you to make more use of the application/device?

What features about the application did you like in its current implementation?

Did you understand what the application was showing you?

How accurate do you think the data was? Was it all equally accurate, were there bits that you trusted more than others?

**If not accurate:** Do you think the data could still be useful given those inaccuracies?

Does the accuracy of the data change how you might use this technology?

Do you think the utility/usefulness would change over time with additional data collection?

How confident would you be making decisions based on the data you were presented with?

What would have made you more confident?

Which features of the Cognizance Tracker application do you feel you make the most use of?

Do you think the amount of information shown in the app was too much, just right, too little?

**Too much:**

What would you like to see less of? / What you like it to focus on?

**Too little:**

What would you like to see more of?

**Just right:**

Last Updated: 30/09/2019



Is there anything you would have liked to have seen more or less of?

Did you trust the values that were being shown? Why (not)?  
Were there any more than others?

**Previous PI exp. only**

How did your usage of the Cognizance Tracker application compare to your previous experience with tracking applications e.g. [application used as per pre-screen]

Which parts of the application did you find most useful?  
Why?

**Insight Feed**

Now I would like to discuss the 'Insight Feed' section of the Cognizance application.  
Did you make use of this section?

*If used "but not much" or response unclear ask all questions.*

**Yes**

How did you use it?

Which types of items did you find least and most...

Interesting?

Trustworthy?

Useful?

Understandable?

Were there any metrics that you thought you'd find but didn't?

Were there any metrics that you thought wouldn't be there but were?

When were you most likely to use the Insight Feed area?

**No**

Why not?

What would have encouraged you to make more use of this section?

**Query Area**

Now I would like to discuss the 'query area' of the Cognizance Tracker application...  
Did you make use of this section?



*If used “but not much” or response unclear ask all questions.*

**Yes**

How did you use it?

Which types of query did you use most?

Were you able to find the data you were looking for?

How did you then use the data you found?

Were there any types of query you wanted to do but were unable to?  
What did you do? (e.g. self-track, workaround, leave).

When did you use the Query Area?

**No**

Why not?

What would have encouraged you to make more use of this section?

**Comparison**

In what situations do you think it was better to use the Query Area instead of the Insight Feed?

And when would you use the Insight Feed instead of the Query Area?

Overall, did you prefer to use the Query Area, the Insight Feed or the Overview area?  
Why?

**CFS/ME Specific Questions Only**

Have you previously used tools for tracking data about yourself to manage your symptoms of CFS/ME?

**Yes:**

Which applications/devices?

Do you still use them?

**No:**

Why not?

What did you like about them?

What did you dislike about them?



How did the Cognizance Tracker application compare to the previous tools?

**No:**

Why not?

Would you now?

Can you think of any other data sources (this may be manually or automatically collected) that you think you would like to combine to enable you to better manage your symptoms?

### **3. FUTURE**

*Thinking about the future...*

Can you think of any other data sources that you think you would like to combine to enable you to learn more about yourself?

What features would suggest are added to future implementations?

## Appendix E

### Study 4

## E.1 Information Sheet

## **Participant Information Sheet – Visualisation of Novel Sensor Data in Personal Informatics**

### **Who am I?**

My name is Cillian Dudley and I am a PhD researcher in the Department of Computer Science at the University of Bath. I am conducting this study as part of my PhD research into how people make sense of data from novel sensors used for self-tracking and/or behaviour change.

### **What is this study about?**

The purpose of this study is to gain a better understanding about how the visualisation of data from near-future sensing technologies impacts peoples' interpretation of these technologies.

### **What will I be asked to do?**

If you decide to take part you will be asked to use two pieces of technology, one that is capable of recording your heart rate that is worn on the wrist, and one that is worn on the head with a sensor placed on your forehead, which allows the recording of brainwaves.

You will be asked to respond to some demographic questions and respond to some initial questions about the technologies you're using.

You will then be asked to take a computer-aided test known as the Stroop test. During this test words will be displayed on screen and you must correctly respond with the colour of text displayed, and not the words itself, by pressing a key on a keyboard. As you progress through the levels of the Stroop test the time you have to respond between words will decrease, this is likely to induce frustration and/or mild stress.

You will perform one run-through of the Stroop test and then be asked to provide a response of how you expect the sensors would record your data.

You will then be asked to perform a second iteration of the Stroop test. Afterwards you will be presented with visualisations of the data recorded by the sensors during the second Stroop test and you will be asked to answer some scaled questions and provided answers to some open-ended questions about the visualisation of the data that was recorded during the second Stroop test.

As a thank you for your participation in this study you will be entered into a draw to win one of three a £20 Amazon gift vouchers.

### **Do I have to take part?**

Taking part in this research is entirely voluntary, and you are free to make your own choice about whether you want to participate. If you agree to take part you can choose not to answer any questions that you do not want to and you are free to withdraw at any time.

You have the right to have your questions about the procedures answered. If you have any questions as a result of reading this information sheet, you should ask the researcher before the study begins.

### **What will happen to the information I provide?**

All data recorded by the sensors will not be identifiable and will only be stored to enable reproduction of the visualisations that were shown to you. Survey responses will also be stored in an anonymized format and will not be directly identifiable by them. Your data will be identified only by a participant number or a pseudonym not associated with your identity.

Once the project is completed, the information you have given to me will be kept safely by the University of Bath. If you give your consent, it may be used by other genuine researchers, with the University of Bath's approval, under PREC NOV 2019

the strict rules governing the confidentiality of your information. So again, your name, or any material that might identify you, will never be used or given to anyone.

**What will happen to the results of this research?**

Survey results will be analysed quantitatively and I may use extracts taken from what you have told me in open-ended responses, however these would not identify you to anyone. The findings of the research may also be published in research journals or used in presentations. If you would like to be sent a summary of the findings, we can arrange for this.

**University of Bath privacy notice**

The University of Bath privacy notice can be found here: <https://www.bath.ac.uk/corporate-information/university-of-bath-privacy-notice-for-research-participants/>.

**What do I do if I would like to take part or have any more questions?**

You can contact me, Cillian Dudley, to arrange a suitable time or to discuss any questions you might have.

Email – [C.Dudley@bath.ac.uk](mailto:C.Dudley@bath.ac.uk)

You can also speak to the supervisor of the project, Dr Simon L. Jones.

Email – [S.L.Jones@bath.ac.uk](mailto:S.L.Jones@bath.ac.uk) Phone – +44 (0)12 2538 5927

If you have any concerns related to your participation in this study please direct them to the Chair of the Department of Psychology Research Ethics Committee, email: [psychology-ethics@bath.ac.uk](mailto:psychology-ethics@bath.ac.uk).  
Our address is

Department of Psychology,  
University of Bath,  
Claverton Down,  
Bath,  
BA2 7AY

PREC reference number: 20-011

Many thanks for taking the time to read this. I would be delighted if you would be willing to take part.

## E.2 Consent Form

**CONSENT FORM**  
**Visualisation of Novel Sensor Data in Personal Informatics**

Please answer the following questions to the best of your knowledge

	YES	NO
<b>DO YOU CONFIRM THAT YOU:</b>		
• are at least 18 years of age?	<input type="checkbox"/>	<input type="checkbox"/>
<b>HAVE YOU:</b>		
• been given information explaining about the study?	<input type="checkbox"/>	<input type="checkbox"/>
• had an opportunity to ask questions and discuss this study?	<input type="checkbox"/>	<input type="checkbox"/>
• received satisfactory answers to all questions you asked?	<input type="checkbox"/>	<input type="checkbox"/>
• received enough information about the study for you to make a decision about your participation?	<input type="checkbox"/>	<input type="checkbox"/>
<b>DO YOU UNDERSTAND:</b>		
• that you are free to withdraw from the study and free to withdraw your data prior to anonymization		
• at any time?	<input type="checkbox"/>	<input type="checkbox"/>
• without having to give a reason for withdrawing?	<input type="checkbox"/>	<input type="checkbox"/>

**I hereby fully and freely consent to my participation in this study**

I understand the nature and purpose of the procedures involved in this study. These have been communicated to me on the information sheet accompanying this form.

I understand and acknowledge that the investigation is designed to promote scientific knowledge and that the University of Bath will use the data I provide for no purpose other than research.

I understand the data I provide will be kept **confidential**. My name or other identifying information will not be disclosed in any presentation or publication of the research.

I understand that the University of Bath may use the data collected for this project in a future research project but that the conditions on this form under which I have provided the data will still apply.

Participant's signature: \_\_\_\_\_ Date: \_\_\_\_\_

Name in BLOCK Letters: \_\_\_\_\_

**Final consent**  
**Having participated in this study**

I agree to the University of Bath keeping and processing the data I have provided during the course of this study. I understand that these data will be used only for the purpose(s) set out in the information sheet, and my consent is conditional upon the University complying with its duties and obligations under the Data Protection Regulation.

Participant's signature: \_\_\_\_\_ Date: \_\_\_\_\_

Name in BLOCK Letters: \_\_\_\_\_

If you have any concerns related to your participation in this study please direct them to the Department of Psychology Research Ethics Committee, via email: [psychology-ethics@bath.ac.uk](mailto:psychology-ethics@bath.ac.uk).



## E.3 Electronic Survey Questions

Visualisation of Novel Sensor Data in Personal Informatics

**Demographic Questions**

Please provide the following demographic information, which will be used for statistical purposes only.

**Q1. Age\***

**Q2. Gender\***

**Q3. Highest Level Of Education\***

☐ University Higher Degree (e.g. MSc, PhD)

☐ First degree level qualification including foundation degrees, graduate membership of a professional Institute, PGCE

☐ A/AS Level (or equivalent)

☐ GCSE/O Level (ore equivalent)

☐ Other school (inc. school leaving exam certificate or matriculation)

☐ None of the above

**Q4. Previous experience with self-tracking/personal informatics\***

☐ None

☐ < 1 month

☐ < 3 months

☐ < 6 months

☐ < 12 months

☐ 1-2 years

☐ 2-5 years

☐ 5+ years

**Q5. Previous experience with brain-computer interfaces\***

☐ None

☐ < 1 month

☐ < 3 months

☐ < 6 months

☐ < 12 months

☐ 1-2 years

☐ 2-5 years

☐ 5+ years

Next

Figure E-1: Page 1

#### Fitbit Versa

Live your best life with Fitbit Versa family—health and fitness smartwatches that last 4+ days and feature activity, sleep and heart rate tracking, notifications, apps, & more.

#### PurePulse

Photoplethysmography is a long word with a short explanation: light can be used to measure blood flow. When your heart beats, blood flows, and the volume of the blood in your wrist changes. Blood—interestingly enough—absorbs green light. The higher your blood volume is, the more green light is absorbed. To calculate blood flow, PurePulse shines a green light onto the skin and uses light detectors called photodiodes to measure how much light is being absorbed. This measurement is used to determine how many times your heart beats per minute.

On a scale of 1 to 10 where one means not at all and 10 means extremely:

**Q6. How would you rate your knowledge of Fitbit Versa device?\***

Not at All Knowledgeable 1 2 3 4 5 6 7 8 9 10 Extremely Knowledgeable

**Q7(a). How confident are you that this device will reflect your heart rate?\***

Not at All Confident 1 2 3 4 5 6 7 8 9 10 Extremely Confident

**Q7(b). Why?\***

Previous

Next

Figure E-2: Page 2

### NeuroSky Mindwave

The MindWave Mobile 2 safely measures and outputs the EEG power spectrums (alpha waves, beta waves, etc), NeuroSky eSense meters (attention and meditation) and eye blinks. The device consists of a headset, an ear-clip, and a sensor arm. The headset's reference and ground electrodes are on the ear clip and the EEG electrode is on the sensor arm, resting on the forehead above the eye (FP1 position). It uses a single AAA battery with 8 hours of battery life.

#### Attention

The Attention Meter algorithm indicates the intensity of mental "focus" or "attention." The value ranges from 0 to 100. The attention level increases when a user focuses on a single thought or an external object, and decreases when distracted. Users can observe their ability to concentrate using the algorithm. In educational settings, attention to lesson plans can be tracked to measure their effectiveness in engaging students. In gaming, attention has been used to create "push" control over virtual objects.

#### Meditation

The Meditation Meter algorithm indicates the level of mental "calmness" or "relaxation." The value ranges from 0 to 100, and increases when users relax the mind and decreases when they are uneasy or stressed. The Meditation Meter quantifies the ability to find an inner state of mindfulness, and can thus help users learn how to self correct and find inner balance to overcome the stresses of everyday life. The algorithm is also used in a variety of game-design controls.

On a scale of 1 to 10 where one means not at all and 10 means extremely:

**Q8. How would you rate your knowledge of the 'NeuroSky MindWave' device?\***

Not At All Knowledgeable 1 2 3 4 5 6 7 8 9 10 Extremely Knowledgeable

**Q9(a). How confident are you that this device will reflect your level of Attention?\***

Not At All Confident 1 2 3 4 5 6 7 8 9 10 Extremely Confident

**Q9(b). Why?\***

**Q10(a). How confident are you that this device will reflect your level of Meditation?\***

Not At All Confident 1 2 3 4 5 6 7 8 9 10 Extremely Confident

**Q10(b). Why?\***

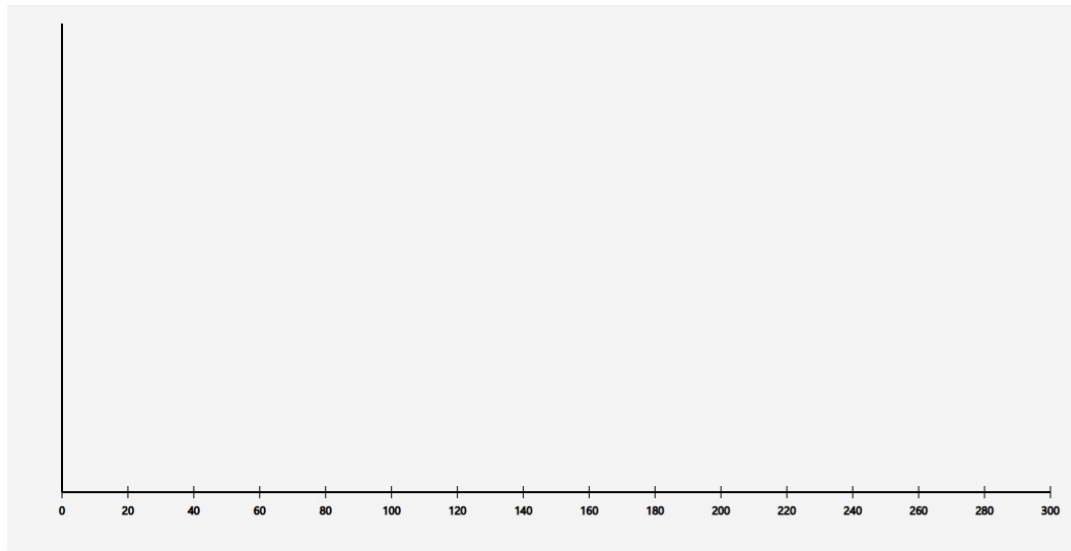
Previous

Complete

Figure E-3: Page 3

## Heart Rate

On a scale of 1 to 10 where one means not at all and 10 means extremely:



Q11. How confident are you that the data shown here accurately reflects your heart rate during the task?\*

Not at All Confident           Extremely Confident

Q12. How closely does the data shown here reflect your prior expectation about how your heart rate might change during the task?\*

Not at All Closely           Extremely Closely

Q13(a). How confident are you that the device used to collect this data is trustworthy?\*

Not at All Confident           Extremely Confident

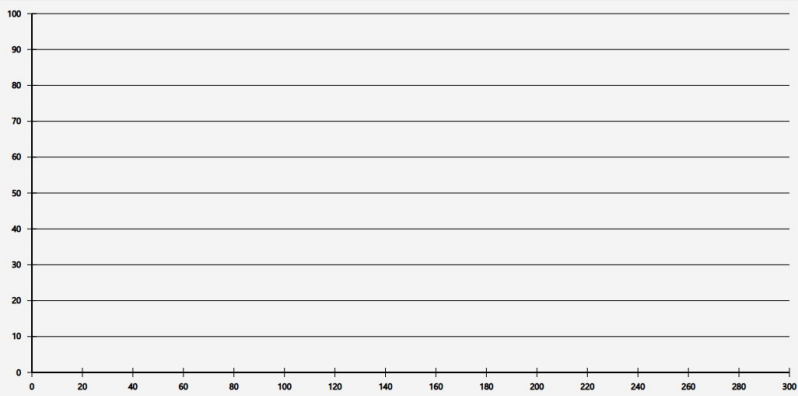
Q13(b). Why?\*

Next

Figure E-4: Page 4

Attention

On a scale of 1 to 10 where one means not at all and 10 means extremely:



Q14. How confident are you that the data shown here accurately reflects your level of Attention during the task?\*

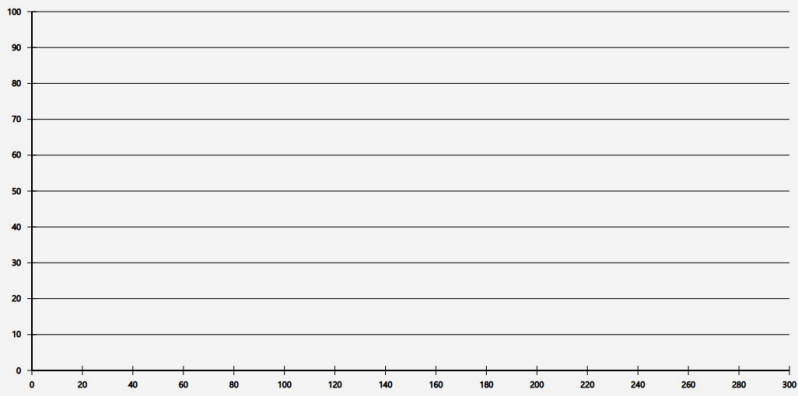
Not at All Confident 1 2 3 4 5 6 7 8 9 10 Extremely Confident

Q15. How closely does the data shown here reflect your prior expectation about how your level of Attention might change during the task?\*

Not at All Closely 1 2 3 4 5 6 7 8 9 10 Extremely Closely

Meditation

On a scale of 1 to 10 where one means not at all and 10 means extremely:



Q16. How confident are you that the data shown here accurately reflects your level of Meditation during the task?\*

Not at All Confident 1 2 3 4 5 6 7 8 9 10 Extremely Confident

Q17. How closely does the data shown here reflect your prior expectation about how your level of Meditation might change during the task?\*

Not at All Closely 1 2 3 4 5 6 7 8 9 10 Extremely Closely

Q18(a). How confident are you that the device used to collect this data is trustworthy?\*

Not at All Confident 1 2 3 4 5 6 7 8 9 10 Extremely Confident

Q18(b). Why?\*

Previous

Complete

Figure E-5: Page 5  
334

## Appendix F

# Data Management Plans

## F.1 General

## Postgraduate Data Management Plan

### 1 Overview

<b>1.1 Project name</b>	Personal Informatics and the Integration of Data from Emerging Sensing Technologies
<b>1.2 Plan author</b>	Cillian Dudley

#### 1.3 Project description

Research undertaken as part of this PhD will provide novel algorithms, applications, visualisations and gain insights regarding the integration of novel sensor data into systems used to reflect on and analyse data for the purpose of self-improvement or -understanding. In particular the research will have a focus on integrating brain activity data (EEG data).

### 2 Compliance

#### 2.1 With what legislative, contractual and policy requirements must the project comply?

Informed consent must be obtained from participants for data to be retained, shared, and used for new purposes.

Access to personal data must be restricted to myself and my supervisor.

The data underlying published results must be kept for at least ten years.

Sources: University of Bath Research Data Policy (<http://www.bath.ac.uk/research/data/policy/>)

### 3 Gathering data

#### 3.1 What data will the project require?

Copies of consent forms will be stored in one ring binder. Qualitative interviews will be recorded and will then be transcribed and analysed with NVivo. User-generated data will be recorded as a result of participants using mobile or web applications developed as part of the project. Quantitative analysis of user application usage may be performed. I expect to generate 2-3GB of data in total.



### 3.2 How will these data be gathered?

Consent forms will be printed and signed by participants. Interviews will be recorded digitally and transcribed into text. User-generated data will be gathered via mobile and web applications interfaces. This data will then be processed in various ways to produce visualisations for the users. Usage logs will be gathered by users when using applications or when interacting with generated visualisations.

### 3.3 What original software, if any, will the project create?

An Android application to allow the recording of EEG data from consumer-grade BCI headset., as well as web and mobile interfaces to allow users to interact and evaluate their data.

## 4 Working with data

### 4.1 Where and how will the data be stored?

My primary copies will be stored on the University's managed data storage (the X Drive), to which both my supervisor and I have access. When working away from a secure and reliable network connection, I will synchronise the files I need between the X Drive and my local hard drive beforehand and immediately afterwards.

User-generated data will be stored in a secured database online database with sufficient backup capabilities. Database dumps will be stored, as required, securely on the X drive.

Code generated as part of the project will be stored on the University's GitHub version control system.

### 4.2 How will access be controlled?

Only my supervisor and I will have access to my data during the project. We will have the only copies of the key to the locked filing cabinet, and the decryption password for encrypted folders.

Access to the user-generated data will be available to the researchers via secure password and users will only be able to access their own data, which will be secured by a personal password of their choosing. These access passwords will be stored in an encrypted format.

Code stored on the University's GitHub account access is restricted to members of the University, where deemed appropriate access will be further restricted such that the repositories are private. Editing of the code will be restricted to collaborators only.

### 4.3 How will the data be organised?

Each individual study will be kept in a separate folder which notes the study name. Inside these transcribed interviews will be named by their transcription data and assigned participant numbers.

Database dumps will be stored as the database name followed by the date of the dump in YYYYMMDD format e.g. CogTracker\_20170817.sql

I will use a Git repository hosted on the University's GitHub to manage the code that I write.

Binaries which are superseded will be labelled with their version number, for quick reference.

#### **4.4 What documentation will accompany the data?**

Any notes taken regarding qualitative data will be provided within or in addition to the NVivo workspace file.

Final code used in the creation of applications used within the project will be commented sufficiently to understand the processes that are undertaken.

## **5 Archiving data**

### **5.1 Which data should be retained long-term? Which will be deleted at the end of the project?**

Anonymised transcripts of all interviews can be kept, but the original audio recordings will be destroyed in order to remove the risk of accidental disclosure.

Backups of user-generated datasets may be useful for future studies and will be archived as database dump files (after removing any identifying features).

Anonymous processed data can be retained if thought to be beneficial to future research, otherwise it should be destroyed.

### **5.2 How will retained data be preserved? For how long?**

I will publish my data in the University's Research Data Archive, where it will be kept for a minimum of ten years.

### **5.3 How will any original software be maintained after the project?**

## **6 Sharing data**

### **6.1 Will access be restricted to any retained data? Why, and how?**

Anonymized summary data will be shared openly.

User-generated data will not be made available to unless informed consent has been obtained and anonymity can be assured, which may mean removing any identifying features before making the dataset available.

## 7 Implementation

### 7.1 How will this plan be kept up to date?

My supervisor and I will review this plan every 6 months and will agree updates if necessary.

### 7.2 What special resources will this plan require, if any?

n/a

### 7.3 What training or further information will you need, if any?

n/a

## F.2 Study III

## 1 Overview

### 1.1 Project title

Understanding how people living with Chronic Fatigue/Myalgic Encephalomyelitis utilise data from emerging sensing technologies for Personal Informatics tools.

### 1.2 Student name and department

Cillian Dudley, Department of Computer Science

### 1.3 Supervisor(s)

Note: the main University of Bath supervisor is the [Data Steward](#) for the project.

**Lead Supervisor:** Dr. Simon L. Jones

### 1.4 Project description

The aim of this research is to explore how people living with Chronic Fatigue Syndrome/Myalgic Encephalomyelitis (CFS/ME) utilise, explore and analyse data from novel sensor technologies. Specifically users will be provided with three pieces of sensing hardware (Fitbit, Spire and a NeuroSky headset) and a smartphone application which acts as a central point to record, access, and review the analysed data. The research will make use of in-the-wild, qualitative and quantitative methods as a means of answering its research questions.

## 2 Compliance

### 2.1 University policy requirements

- All of my research data must be stored securely and backed up (I will use the University's H:/ or X:/ drive)
- Access to the user-generated data will be available to the researchers via secure password and users will only be able to access their own data, which will be secured by a personal password of their choosing. These access passwords will be stored in an encrypted format.
- Informed consent must be obtained from participants for data to be retained, shared, and used for new purposes.
- Data underpinning publications must be kept for at least ten years.
- Published papers must include a data access statement.

#### University policy or guidance

[University of Bath Research Data Policy](#)

[University of Bath Code of Good Practice in Research Integrity](#)

[University of Bath Electronic Information Systems Security Policy](#)

[University of Bath Intellectual Property Policy](#)

[University of Bath Code of Ethics](#)

### 2.2 Legal requirements

All processing of personal data will be carried out in line with the Data Protection Act and GDPR.

#### UK Legislation or framework

*Data Protection Act*

GDPR	
<b>2.3 Contractual requirements</b>	
Not Applicable	
<b>Name of funder</b>	<b>Data policy URL</b>

### 3 Gathering data

<b>3.1 Description of the data</b>
<b>3.1.1 Types of data</b>
<ul style="list-style-type: none"> <li>• I will be using data taken from the published literature.</li> <li>• I will be generating qualitative data from interviews or meetings with participants.</li> <li>• I will be generating qualitative data from user testing of prototypes.</li> <li>• I will be generating quantitative data from user testing of prototypes.</li> <li>• I will be writing software code.</li> </ul>
<b>3.1.2 Format and scale of the data</b>
<ul style="list-style-type: none"> <li>• I will generate up to 1 GB of data including audio recordings and user collected data.</li> <li>• I will use Word (.docx) to store text documents, such as questionnaires, study plans and participant contact sheets.</li> <li>• I will use Excel (.xlsx) and comma separated values (.csv) documents to store data from questionnaires, for example.</li> <li>• I will use LaTeX to write publications (.tex and others), which are basic text files.</li> <li>• Images, audio files, and video content will be saved using a recommended or acceptable format.</li> <li>• I expect my consent forms to less than one ring binder.</li> </ul>
<b>3.2 Data collection methods</b>
<ul style="list-style-type: none"> <li>• Consent forms will be printed and signed by participants.</li> <li>• Interviews will be recorded digitally and transcribed into text.</li> <li>• User-generated data will be gathered via mobile and web applications interfaces. This data will then be processed in various ways to produce visualisations for the users.</li> <li>• Usage logs will be gathered by users when using applications or when interacting with generated visualisations.</li> </ul>
<b>3.3 Development of original software</b>
<p>An Android smartphone application to allow the recording of EEG data from consumer-grade BCI headset as well as presenting and allowing the ability to analyse user data has been produced. This application depends on a web server application capable of synching user recorded EEG data between the smartphone application and the server. The server application also requests data from Fitbit and Spire endpoints which are then synchronised back to the mobile application.</p>

### 4 Working with data

<b>4.1 Short- and medium-term data storage arrangements</b>
---

Copies of consent forms will be stored in one ring binder. Qualitative interviews will be recorded and will then be transcribed and analysed with NVivo. User-generated data will be recorded as a result of participants using mobile or web applications developed as part of the project. Quantitative analysis of user application usage may be performed. I expect to generate 2-3GB of data in total.

#### **4.2 Control of access to data and sharing with collaborators**

Only my supervisor and I will have access to my data during the project. We will have the only copies of the key to the locked filing cabinet, and the decryption password for encrypted folders.

Access to the user-generated data will be available to the researchers via secure password and users will only be able to access their own data, which will be secured by a personal password of their choosing. These access passwords will be stored in an encrypted format.

Code stored on the University's GitHub account access is restricted to members of the University, where deemed appropriate access will be further restricted such that the repositories are private. Editing of the code will be restricted to collaborators only.

#### **4.3 File organisation and version control**

This study's research materials will be kept in a separate folder which notes the study name. Inside these transcribed interviews will be named by their transcription data and assigned participant numbers. Transcriptions will be kept with audio until such time as they are to be deleted.

Database dumps will be stored as the database name followed by the date of the dump in YYYYMMDD format e.g. CogTracker\_20170817.sql

I will use a Git repository hosted on the University's GitHub to manage the code that I write. Binaries which are superseded will be labelled with their version number, for quick reference.

#### **4.4 Documentation that will accompany the data**

Any notes taken regarding qualitative data will be provided within or in addition to the NVivo workspace file.

Final code used in the creation of applications used within the project will be commented sufficiently to understand the processes that are undertaken.

## **5 Archiving data**

### **5.1 Selection of data to be retained and deleted at the end of the project**

Anonymised transcripts of all interviews can be kept, but the original audio recordings will be destroyed in order to remove the risk of accidental disclosure.

Backups of user-generated datasets may be useful for future studies and will be archived as database dump files (after removing any identifying features).

Anonymous processed data can be retained if thought to be beneficial to future research, otherwise it should be destroyed.

### **5.2 Data preservation strategy and retention period**

I will publish my data in the University's Research Data Archive, where it will be kept for a minimum of ten years.

### **5.3 Maintenance of original software**

A final binary file will be kept archived with the data. Source code will remain available on GitHub. Currently intentions for maintenance are unknown.

## 6 Sharing data

### 6.1 Justification for any restrictions on data sharing

Anonymized summary data will be shared openly.

User-generated data will not be made available to unless informed consent has been obtained and anonymity can be assured, which may mean removing any identifying features before making the dataset available.

### 6.2 Arrangements for data sharing

If and when I must share data (for example, data underpinning a publication), I will share my data openly via the University of Bath Research Data Archive. I will include the DOI to my dataset in data access statements provided in publications from my project.

When appropriate, I will share my code openly through GitHub and will archive a snapshot of the code underpinning relevant publications in Zenodo. The DOI to the Zenodo record, and the URL to the GitHub record will be included in data access statements provided in publications from my project.

## 7 Implementation

### 7.1 Review of the Data Management Plan

My supervisor and I will review this plan every 6 months and will agree updates if necessary.

### 7.2 Special resources required for the project

Not applicable.

### 7.3 Further training needs

Not applicable.



## Appendix G

### Miscellaneous

## G.1 13-point Ethics Checklist

UNIVERSITY OF BATH

**Department of Computer Science**

**13-POINT ETHICS CHECK LIST**

This document describes the 13 issues that need to be considered carefully before students or staff involve other people ("participants") for the collection of information as part of their project or research.

1. *Have you prepared a briefing script for volunteers?*  
You must explain to people what they will be required to do, the kind of data you will be collecting from them and how it will be used.
2. *Will the participants be using any non-standard hardware?*  
Participants should not be exposed to any risks associated with the use of non-standard equipment: anything other than pen and paper or typical interaction with PCs on desks is considered non-standard.
3. *Is there any intentional deception of the participants?*  
Withholding information or misleading participants is unacceptable if participants are likely to object or show unease when debriefed.
4. *How will participants voluntarily give consent?*  
If the results of the evaluation are likely to be used beyond the term of the project (for example, the software is to be deployed, or the data is to be published), then signed consent is necessary. A separate consent form should be signed by each participant.
5. *Will the participants be exposed to any risks greater than those encountered in their normal work life?*  
Investigators have a responsibility to protect participants from physical and mental harm during the investigation. The risk of harm must be no greater than in ordinary life.
6. *Are you offering any incentive to the participants?*  
The payment of participants must not be used to induce them to risk harm beyond that which they risk without payment in their normal lifestyle.
7. *Are any of your participants under the age of 16?*  
Parental consent is required for participants under the age of 16.
8. *Do any of your participants have an impairment that will limit their understanding or communication?*  
Additional consent is required for participants with impairments.

9. *Are you in a position of authority or influence over any of your participants?*

A position of authority or influence over any participant must not be allowed to pressurise participants to take part in, or remain in, any experiment.

10. *Will the participants be informed that they could withdraw at any time?*

All participants have the right to withdraw at any time during the investigation. They should be told this in the introductory script.

11. *Will the participants be informed of your contact details?*

All participants must be able to contact the investigator after the investigation. They should be given the details of the Unit Lecturer or Supervisor as part of the debriefing.

12. *Will participants be de-briefed?*

The student must provide the participants with sufficient information in the debriefing to enable them to understand the nature of the investigation.

13. *Will the data collected from the participants be stored in an anonymous form?*

All participant data (hard copy and soft copy) should be stored securely, and in anonymous form.

**NAME:** \_\_\_\_\_

**SUPERVISOR (IF APPLICABLE):** \_\_\_\_\_

**SECOND READER (IF APPLICABLE):** \_\_\_\_\_

**PROJECT TITLE:** \_\_\_\_\_

**DATE:** \_\_\_\_\_